

ZAPATA SAPIENCIA, DANIELA. Ph.D. Essays on Health Insurance Coverage and Food Assistance Programs (2012).
Directed by Dr. David Ribar. 143 pp.

Empirical work shows that health insurance coverage improves children's health and that healthier children have better educational and labor market outcomes. This suggests that the benefits of higher insurance rates among children go beyond improvements in health. However, there are no investigations in the United States that track the long-term socioeconomic benefits of health insurance coverage during childhood. Using data from the Children of the National Longitudinal Survey of Youth to estimate family fixed effects models, I find evidence that health insurance coverage at ages 0-4 has a positive effect on test scores in mathematics, reading recognition, reading comprehension, and vocabulary at ages 5-14.

The second essay in this dissertation, co-authored with Charles Courtemanche, investigates the effect of the Massachusetts health care reform on self-reported health. The main objective of this reform was to achieve universal health insurance coverage through a combination of insurance market reforms, mandates, and subsidies. This reform was later used as a model for the Patient Protection and Affordable Care Act (ACA). Using individual-level data from the Behavioral Risk Factor Surveillance System and a difference in differences estimation strategy, this essay provides evidence that this reform led to better overall self-assessed health. Several determinants of overall health, including physical health, mental health, functional limitations, joint disorders, body mass index, and moderate physical activity also improved.

Public food assistance programs share the fundamental goal of helping needy and vulnerable people in the U.S. obtain access to nutritious foods that they might not otherwise be able to afford. These programs also have other objectives, such as improving recipients' health, furthering children's development and school performance. To investigate these broader impacts, the third chapter of this dissertation, co-authored with David Ribar, examines the relationship between participation in food assistance programs, family routines and time use. Results from fixed effects models estimated using longitudinal data from the Three-City Study indicate that SNAP participation is negatively associated with homework routines. WIC participation on the other hand, is positively associated with family routines in general and with dinner routines, homework routines, and family-time routines in particular.

ESSAYS ON HEALTH INSURANCE COVERAGE AND FOOD ASSISTANCE
PROGRAMS

by

Daniela Zapata Sapiencia

A Dissertation Submitted to
the Faculty of the Graduate School at
The University of North Carolina Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Greensboro
2012

Approved by

Dr. David Ribar
Committee Chair

© 2012 Daniela Zapata Sapiencia

APPROVAL PAGE

This dissertation has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair _____
David Ribar

Committee Members _____
Christopher Ruhm

Charles Courtemanche

Christopher Swann

August 27, 2012
Date of Acceptance by Committee

August 27, 2012
Date of Final Oral Examination

ACKNOWLEDGEMENTS

I want to acknowledge my advisor, David Ribar, for his guidance, generous attention to my work, and his useful suggestions, which have made this research better. The other members of my Dissertation Committee, Christopher Ruhm, Charles Courtemanche, and Christopher Swann have been indispensable to my progress. In particular, I want to express my deep appreciation for Christopher Ruhm, who patiently guided me in the first essay of this dissertation, and was fundamental in my improvement as a researcher. The mentorship and support of Charles Courtemanche push me to learn more and have been invaluable in my academic growth, for which I am extremely grateful.

I would like to thank to Ken Snowden and Stephen Holland for their friendly advice and general collegiality that they offered me over the years. I also want to acknowledge my friend and mentor Diana Kruger for her guidance and advice, which encourage me to pursue this Doctoral degree.

This work would have been impossible without the loving support of my husband Marcelo Ochoa.

I thank my parents, brother, and sister for their unconditional love.

TABLE OF CONTENTS

	Page
LIST OF TABLES	vi
LIST OF FIGURES	viii
 CHAPTER	
I. HEALTH INSURANCE COVERAGE AND CHILDREN’S COGNITIVE OUTCOMES.....	1
Abstract	1
Introduction	1
Empirical Approach	5
Data	9
Results from Multivariate Models	15
Health Insurance and Children’s Cognitive Outcomes	15
Robustness Checks	18
Heterogeneity	19
Discussion	20
Tables and Figures	23
 II. DOES UNIVERSAL COVERAGE IMPROVE HEALTH? THE MASSACHUSETTS EXPERIENCE	 32
Abstract	32
Introduction	33
Health Insurance and Health	38
Data	41
Regression Analysis	47
Baseline Model	47
Robustness Checks	53
Testing for Differential Pre-Treatment Trends and Delayed Effects	56
Testing for Endogenous Moving Patterns	58
Tests Related to Inference	59
Other Health Outcomes	60
Heterogeneity	64
Instrumental Variables	66
Conclusion	70
Tables and Figures	74

III. FOOD ASSISTANCE AND FAMILY ROUTINES IN THREE AMERICAN CITIES	88
Abstract	88
Introduction	89
Conceptual Approach	92
Data	97
Descriptive Analysis	101
Multivariate Analyses	103
Ordinary Least Squares Results	105
Longitudinal Fixed-Effect Results	106
Sensitivity Analyses	108
Discussion	108
Tables and Figures	111
REFERENCES	115
APPENDIX A. ESTIMATING EFFECT ON HEALTH USING SERIES OF PROBITS	129
APPENDIX B. ESTIMATING EFFECT ON HEALTH USING SERIES OF LINEAR PROBABILITY MODELS	130
APPENDIX C. FALSIFICATION TESTS USING PRE-TREATMENT DATA	131
APPENDIX D. CORRELATIONS BETWEEN OVERALL HEALTH AND OTHER HEALTH OUTCOMES	132
APPENDIX E. INSTRUMENTAL VARIABLES: STRATIFIED BY GENDER AND AGE	133
APPENDIX F. INSTRUMENTAL VARIABLES: STRATIFIED BY AGE AND INCOME	135
APPENDIX G. ESTIMATES FROM ORDINARY LEAST SQUARES MODELS. FULL RESULTS	136
APPENDIX H. ESTIMATES FROM FIXED EFFECTS MODELS. FULL RESULTS	139

LIST OF TABLES

	Page
Table 1.1. Distribution of Children According to the Proportion of Time they were Covered by Health Insurance	23
Table 1.2. Means of Outcomes and Observable Characteristics According to Health Insurance Coverage during Early Childhood (Ages 0-4).....	24
Table 1.3. Number of Families and Children with Variation in Health Insurance Coverage at Ages 0-4	26
Table 1.4. The Effect of Early Childhood Health Insurance Coverage on Children's Test Scores Taken at Ages 5-14.....	27
Table 1.5. The Effect of Private Health Insurance and Medicaid Coverage on Children's Test Scores Taken at Ages 5-14.....	28
Table 1.6. The Effect of Health Insurance Coverage on Children's Test Scores	29
Table 1.7. The Effect of Health Insurance on Children's Test Scores Taken at Ages 5-14	30
Table 1.8. The Effect of Health Insurance on Children's Test Scores Taken at Ages 5-14	31
Table 2.1. Similarities and Differences between the Massachusetts Reform and the National Reform (ACA).....	75
Table 2.2. Pre-Treatment Means of Health Variables	77
Table 2.3. Pre-Treatment Means of Control Variables.....	78
Table 2.4. Difference-in-Differences Ordered Probit Regressions.....	79
Table 2.5. Testing for Differential Pre-Treatment Trends and Delayed Effects	81
Table 2.6. Testing for Endogenous Moving Patterns	82
Table 2.7. Regressions with Aggregated Data.....	83
Table 2.8. Regression Results for Other Health Outcomes	84

Table 2.9. Heterogeneity in the Effect on Health by Gender and Age	85
Table 2.10. Heterogeneity in the Effect on Health by Race and Income.....	86
Table 2.11. Instrumental Variables	87
Table 3.1. Means of Outcomes and Independent Variables	111
Table 3.2. Selected Estimates from Ordinary Least Squares Models of Family Routines and Daily Schedules	113
Table 3.3. Selected Estimates from Fixed Effects Models of Family Routines and Daily Schedules	114

LIST OF FIGURES

	Page
Figure 2.1. Changes in Health Status Index 2001-2010	74

CHAPTER I

HEALTH INSURANCE COVERAGE AND CHILDREN'S COGNITIVE OUTCOMES

Abstract

Using data from the Children of the National Longitudinal Survey of Youth, this paper finds evidence that health insurance coverage at ages 0-4 has a positive effect on test scores in mathematics, reading recognition, reading comprehension, and vocabulary at ages 5-14. The observation that children without health insurance have worse health than their insured counterparts is one of the motivations behind the efforts to increase health insurance coverage for children, including the recent Patient Protection and Affordable Care Act. This paper provides evidence that the benefits of insurance coverage for children go beyond improvements in health and include lasting effects on children's cognitive outcomes.

Introduction

The large number of people without health insurance coverage and the difficulties they have obtaining adequate medical care are concerns behind the health insurance market reforms over the last two decades, including the recent Patient Protection and

Affordable Care Act. Empirical work shows that in the case of children health insurance coverage improves health (Currie and Gruber, 1996a; 1996b; Currie et al., 2008), and that healthier children have better educational and labor market outcomes (Case, et al., 2005; Currie et al., 2010). This suggests that the benefits of higher insurance rates among children go beyond improvements in health. However, as pointed out by the Baker Institute (2009), there are no investigations in the United States that track the long-term socioeconomic benefits of health insurance coverage during childhood. This paper takes a first step toward filling this void by using a dataset that follows children over time to investigate the effect that health insurance coverage early in childhood has on cognitive outcomes later in childhood.

According to standard economic theory the production of children's health depends on market goods, like medical care, the initial endowment of health, and time invested in health producing activities (Haveman and Wolfe, 1995; Currie, 2009). Health insurance coverage increases the consumption of medical care because it reduces the price faced by the consumer (Arrow, 1963; Pauly, 1968). If this additional medical care is not redundant, it translates into better health. Previous research demonstrates that having health insurance coverage increases medical care consumption (Dafny and Gruber, 2000; Aizer, 2007) and improves health among children (Currie and Gruber 1996a; Currie et al., 2008).

The literature also documents the existence of a positive relationship between childhood health and future cognitive and educational outcomes. For example, Case and Paxson (2010) use children's height for age as a summary measure for health and find

that it is positively correlated with children's cognitive development and school progress. Case et al. (2005) find that chronic conditions at age 7 reduce the number of proficiency tests passed by age 16. Currie et al. (2010) find that physical conditions at ages 9-13 and 14-18 have negative effects on literacy, math achievement and reduce the likelihood of being in grade 12 by age 17. Currie (2009) and Currie et al. (2010) argue that good health has a positive effect on schooling because it raises children's productivity at school, while poor health impairs cognitive skills. If health is an input in the production of cognitive skills, and if these skills are self-producing, meaning that past period skills are an input in the production of future skills (Cunha and Heckman, 2008), then health will also have a lasting effect on cognitive skills.

The links between health and cognitive outcomes and between health insurance and children's health suggest that it is possible that health insurance coverage affects cognitive and educational outcomes. However, it is also possible that the health improvements caused by health insurance coverage are not large enough to affect children's cognitive outcomes. Furthermore, since health is multidimensional it is possible that health insurance improves health in dimensions that are not relevant for the production of cognitive skills.

Understanding the relationship between health insurance and cognitive outcomes is important because according to the U.S. Census Bureau in 2010 more than 8 million children did not have health insurance of any kind. Even though the implementation of the Affordable Care Act is expected to reduce this number greatly it is still possible that

this law is fully or partially repealed or that during the budgetary process it's financing gets reduced preventing its full implementation.

Two previous studies have looked at the effect of health insurance coverage on children's educational outcomes. Chen and Zhe Jin (2012) find that the 2006 health insurance expansions in rural China had a positive effect on the school enrollment of six-year-old children, but find no effect for other cohorts. Levine and Schanzenbach (2009) investigate the effect of Medicaid and SCHIP expansions on fourth and eighth graders state-average test scores in mathematics and reading. They find that a 50 percentage point increase in Medicaid eligibility at birth increases state-average reading test scores by 0.09 of a standard deviation, but find no effect for test scores in mathematics.

This paper contributes to the literature by using individual-level data from the Children of the National Longitudinal Survey of Youth (CNLSY) to investigate the impact of health insurance coverage during early childhood (ages 0-4) on cognitive outcomes measured by standardized test scores in mathematics, reading recognition, reading comprehension, and vocabulary taken during late childhood (ages 5-14). Family and child characteristics are included in regression analyses, and family fixed effects models are used as the main estimation strategy to control for time invariant unobservable characteristics of the household. The paper also investigates if the associations found depend on whether children have private health insurance or Medicaid coverage. These two types of coverage may have different impacts on cognitive outcomes because the intensity and availability of medical care may differ depending on the type of coverage children have. For example, many doctors do not participate in

Medicaid, which increases the waiting times for children seeking medical attention (Skaggs et al., 2006; Hwang et al., 2005). There is also evidence that shows that doctors spend less time with Medicaid patients than with private insured patients (Decker, 2007). In contrast, Currie and Thomas (1995) show that children covered by Medicaid have more doctor checkups than private insured children. Dissimilarities in the medical care received may translate into differences in health outcomes, which may ultimately affect children's cognitive outcomes.

The empirical estimates show that health insurance coverage at ages 0-4 improves Peabody Individual Achievement Test (PIAT) scores in mathematics, reading recognition, reading comprehension and vocabulary at ages 5-14. Both private health insurance and Medicaid coverage have similar effects on these cognitive outcomes. Additional analyses reveal that the effect of health insurance on test scores does not disappear over time. Health insurance coverage at ages 0-4 has a positive effect on PIATs taken at ages 5-9 and on PIATs taken at ages 10-14. Health insurance coverage at ages 5-9 also improves test scores at ages 10-14. Health insurance coverage has a larger effect for boys than for girls, and the evidence suggests that the effects are larger for children living in low-income households.

Empirical Approach

Suppose that the lifetime of a child up to age 14 can be divided in two periods. Early childhood, which starts at the time the child is born and ends at age 4, and late childhood, which begins when the child is 5 and ends at age 14. According to Almond and Currie (2010), investments before the age of 5 have lasting consequences in

children's outcomes and should be distinguished from later investments. The scores of tests taken at late childhood are measures of cognitive achievement; they are produced with contemporaneous inputs that include home investments, health and cognitive skills. Current health depends on the investments made to preserve it, which include health insurance coverage, and past period health. Cognitive skills depend on past period cognitive skills and are also affected by past period health. Replacing the relationships that describe the production of health and cognitive skills in the test scores equation results in the following reduced form equation:

$$y_{i,f,late} = \delta_0 + \delta_1 h_{i,f,endowment} + \delta_2 c_{i,f,endowment} + \delta_3 HI_{i,f,early} + \delta_4 HI_{i,f,late} + \delta_5 X_{i,f,early} + \delta_6 X_{i,f,late} + \varepsilon_{i,f,early-late} \quad (1.1)$$

Equation (1.1) says that the score in a cognitive ability test ($y_{i,f,late}$) for child i in family f taken at late childhood, is a function of endowments of health ($h_{i,f,endowment}$) and cognitive skills ($c_{i,f,endowment}$), health insurance coverage during early ($HI_{i,f,early}$) and late ($HI_{i,f,late}$) childhood, other observable inputs during early ($X_{i,f,early}$) and late ($X_{i,f,late}$) childhood, and other unobservable characteristics during early and late childhood ($\varepsilon_{i,f,early-late}$).

As mentioned, health insurance might affect cognitive outcomes through its effect on health.¹ If so, the size of the parameters δ_3 and δ_4 depend on the magnitude in which health insurance affects health and the magnitude in which health affects test scores. If health insurance has a large effect on children's health and if the production of test scores

¹ Health insurance coverage could also improve children's cognitive outcomes by other mechanisms such as greater economic stability, reduced stress and earlier identification of cognitive disabilities.

is very sensitive to this input, then we would observe a large effect of health insurance on test scores. If instead health insurance has a positive but small effect on health, or if test scores were not very sensitive to changes in health, then we would observe a small effect of health insurance on test scores.

Estimation of equation (1.1) using Ordinary Least Squares (OLS) yields unbiased estimates of δ_3 and δ_4 when health insurance is not correlated with unobservable characteristics that also affect test scores. Controlling for a detailed list of covariates reduces the possibility of confounders but does not eliminate it. To illustrate this we could re-write the error term in equation (1.1) as: $\varepsilon_{i,f,early-late} = \mu_{i,f,early-late} + \varphi_f$, where φ_f represents time invariant characteristics of the family. Focusing the discussion on the coefficient estimate of early childhood health insurance coverage we have that if $cov(\varphi_f, HI_{i,f,early}) \neq 0$ then the OLS coefficient estimate would be biased. Moreover, if $\varphi_f > 0$ and $cov(\varphi_f, HI_{i,f,early}) > 0$ then the OLS estimate would be upward biased. This could happen if for example φ_f represents family ability and if children in high-ability families were more likely to have higher test scores and also more likely to have health insurance coverage. On the other hand, if $\varphi_f > 0$ and $cov(\varphi_f, HI_{i,f,early}) < 0$ then OLS the estimate would be downward biased. This could happen if, for instance, φ_f represents family's health and if healthier parents were less likely to buy family health insurance coverage and at the same time more likely to have children with higher test scores.

The CNLSY dataset has information on multiple siblings within a household, which allows implementing a family fixed effect estimator. Using this estimation strategy will result in unbiased estimates of the coefficient of interest if there are no remaining

time-varying unobservable family covariates or unobserved individual heterogeneity correlated with health insurance coverage and test scores.

The presence of time-varying unobserved family characteristics associated with health insurance coverage and test scores would bias family fixed effects estimates. Besides including a detailed list of time-varying controls to reduce this probability, I also test the sensitivity of the results to alternative specifications described in more detail in the results section.

The presence of child-specific unobserved heterogeneity would also result in biased fixed effects coefficient estimates. Following Griliches (1979) Bound and Solon (1999) show that the inconsistency of the Fixed Effect (OLS) estimator is proportional to the fraction of between-siblings (cross-sectional) variation in the dependent variable due to the individual endogenous component. For example, if parents exhibit a compensatory behavior and insure only the more disadvantaged children, who are also more likely to have lower test scores, then both FE and OLS would be downward biased, and if parents exhibit a reinforcing behavior and insure only the more talented children then both OLS and FE would be upward biased. However, the inconsistency of the FE estimator would be larger than the OLS estimator if the endogenous variation comprises a larger share of the between-siblings variation in test scores than it does of the cross-sectional variation.²

² Medicaid coefficient estimates could be more sensitive to children unobservables if children obtain coverage only after they suffer a health shock and visit an Emergency Room. However, Medicaid enrollment rates have been above 73% since 1986 (Gruber, 2000; Kenney et al. 2009), which suggests that a large number of children enroll in Medicaid independently of their health status. In contrast, the characteristics of private health insurance policies reduce the probability of enrollment based on child specific cofounders because family policies have fixed enrollment periods, require waiting periods before covering people with pre-existing conditions, and generally charge the same premium regardless of the number of children being covered. Under the HIPAA insurance companies could establish waiting period

The use of family fixed effects can also aggravate measurement error problems relative to OLS (Griliches, 1979). If there is random measurement error of the independent variables of interest OLS coefficient estimates are downward biased. Black et al. (2000) show that when the independent variables of interest are binary, which is the way health insurance coverage is measured, OLS estimates would also be downward biased if the variance of the measurement error is smaller than the covariance between the true value of the variable and its measurement error. Since family fixed effects exacerbate measurement error bias, if there is measurement error in the health insurance variables and if the covariance between the true value of health insurance and its measurement error is smaller than the variance of the measurement error, then fixed effects estimates would represent a lower bound of the true effect.

Data

The National Longitudinal Survey of Youth (NLSY79) is a nationally representative sample of men and women (when weighted) born between January 1, 1957 and December 31, 1964 who have been surveyed since 1979. In 1986, the Children of the National Longitudinal Survey of Youth began interviewing the sons and daughters of female NLSY79 respondents every two years. By 2008, a total of 10,488 children born to NLSY79 female respondents had been surveyed at some point.

The paper's conceptual framework describes the lifetime of a child as a two-period model: early childhood, ages 0 to 4, and late childhood, ages 5 to 14. The main

for preexisting conditions of up to 12 months (up to 18 months for late enrollees). These waiting periods could be shorter with if the person can show prior insurance coverage of that preexisting condition. Before HIPAA was enacted insurers could choose not to cover preexisting conditions.

interest of this paper is to measure the effect of health insurance coverage during early childhood on test scores measured at late childhood. This means that children in the estimating sample have to be interviewed at least once during each of these periods. Furthermore, children have to be interviewed at least once before they were 2 years old to have measures of family and individual conditions, including health insurance coverage, around the time of birth. This means that only those born between 1984 and 2003 can be part of the estimating sample. There are 6,572 children surveyed by the CNLSY that satisfy this restriction, of which 5,017 were first surveyed before age 2 and re-interviewed after age 5.³ Family fixed effects require that children in the estimating sample have siblings who were also surveyed; there are 3,508 children who satisfy this condition. It is important to note that even though late childhood is defined as ages 5-14, children were on average 12 years old the last time they were interviewed and there are some children who were as young as 5 years old the last time they participated on the CNLSY.

This paper uses the scores in the Peabody Individual Achievement Test in mathematics (PIAT-M), reading recognition (PIAT-R), reading comprehension (PIAT-C) and the Peabody Picture Vocabulary Test - Revised (PPVT) to measure cognitive outcomes. The PIAT-M gauges attainment in mathematics, the PIAT-R measures word recognition and pronunciation, the PIAT-C evaluates a child's ability to derive meaning from sentences that are read silently, and the PPVT provides an estimate of verbal ability

³ Not all eligible children are interviewed every survey wave because of attrition. For instance, between 1986 and 1998 the CNLSY could not collect information for 10 to 20 percent of all eligible children (Aughinbaugh, 2004).

by evaluating hearing vocabulary for standard American English. These assessments have been widely used to measure cognitive achievements (see e.g. Cunha et al., 2010; Baum, 2003; Ruhm, 2004; and Ruhm, 2008). PIATs are administered every survey year to all children ages 5 to 14, while the PPVT follows a more complex administration pattern. Every survey wave 10-11 year old children take the PPVT, along with those who previously missed an assessment Children older than 3 were also tested in 1986 and in 1992.

PIATs, and achievement tests in general, measure how much knowledge the test taker has accumulated in a particular area. Unlike intelligence tests, which are set before the age of 10, achievement tests are quite malleable and therefore sensitive to investments made to improve them (Cunha and Heckman, 2008). Cognitive ability, measured by achievement tests like the PIAT, has been found to have a positive effect on educational achievement, employment, and wages (Heckman et al., 2006). Therefore, even though the CNLSY does not have information on children's school grades, or enough information to examine children's future labor market performance, the results found in this paper provide an indication of the potential effect of health insurance coverage on these other outcomes.

The analysis in this paper uses age-specific normed test scores provided by the CNLSY that were re-normalized to have a mean of 10 and a standard deviation of 1 to facilitate the interpretation of the results.⁴ The majority of children were interviewed more than once during this period, which means that they have more than one

⁴ Unreported regressions using the original age-normed scores provided by the CNLSY give results that are qualitatively the same to the ones reported on this paper.

observation for each test. For this reason the dependent variables used in the regression analyses are average test scores obtained during late childhood (ages 5-14). Because cognitive outcomes can change a lot during this time frame, I also perform separate analyses for test scores obtained at ages 5-9 and at ages 10-14.

Every wave of the core survey of the NLSY79 asks mothers two questions about the health insurance status of their children. One of them asks if the health care of the child is covered by an individual or employer-provided plan; the other question asks whether Medicaid covers the health care of the child. Using these measures I create three dummy variables to measure health insurance coverage: one for private health insurance coverage, a second for Medicaid coverage, and a third for health insurance coverage in general, either Medicaid or private health insurance.⁵ Early and late childhood health insurance coverage is measured using the average of these dummy variables over early and late childhood respectively. Separate analyses use averages over ages 0-4, 5-9 and 10-14.

The CNSLY obtains information about household and children's time-varying characteristics. These data include household income, the HOME-SF score (a measure oriented to capture the quality of the home environment), the number and ages of people living in the household, welfare participation, mother's: education, marital, and employment status, employment status of the spouse, and children's age.⁶ The CNSLY

⁵ If the mother reported that the child had both Medicaid and private health insurance, I assume that the child was using private health insurance coverage only.

⁶ The HOME-SF contains self-reported information on time spent by the mother promoting cognitive stimulation of the children, mother emotional support, and home investments that facilitate learning. This measure has been used extensively to capture home inputs in the production of cognitive skills (e.g. Todd and Wolpin, 2007; Ruhm, 2008; Cunha and Heckman, 2008).

also obtains information about household and children's time invariant characteristics such as: age of the mother at the time of birth, number of drinks and cigarettes the mother consumed during pregnancy, which are used as proxy measures of endowments of children's health and cognitive skills, child's fetal growth (i.e. birthweight divided by length of pregnancy), race, birth order and year of birth. Regression analyses also control for the year of the child's birth and the number of times the child was interviewed during early and late childhood. For some children information on one or more control variables is missing. To avoid excluding these observations from the analysis, the relevant regressors are set to zero and dummy variables were created denoting the presence of missing values. For example, if family income was not reported this variable is given a value of zero and the dummy variable "missing family income" is set to one.⁷

Table 1.1 shows health insurance coverage patterns (unweighted) at ages 0-4, 5-9 and 10-14 for children in the sample. The category "always health insurance" coverage means that every time the child was interviewed by the CNLSY the mother reported that he had health insurance coverage (either private or Medicaid). It is possible that the child had no health insurance between surveys, but that information is not collected by the CNLSY. Table 1.1 shows that during early childhood 84.6 percent of children always had health insurance coverage. Private health insurance is the most common form of insurance, with 65.9 percent of children always having that type of coverage during early

⁷ This was done for fetal growth, average family income during early and late childhood, average welfare participation during early and late childhood, average HOME-SF score during early and late childhood, average education, marital, and employment status of the mother during early and late childhood, number of cigarettes and number of drinks consumed during pregnancy.

childhood. Medicaid coverage is less frequent; during early childhood 9.7 percent of children can be classified as always having that type of coverage.

The first column of Table 1.2 shows sample means of test scores taken at ages 5-14 and of observable characteristics during early and late childhood for all the children in the sample. Table 1.2 also shows sample means for children classified as “always with health insurance coverage of any type”, “always with private health insurance”, “always with Medicaid coverage” and “always uninsured” at ages 0-4.⁸ Stars denote that for children with health insurance coverage the number is statistically different from the “always uninsured” mean. The top panel of the Table shows that children in the “always health insurance” category have higher test scores than uninsured children, but if children are further divided according to type of coverage it can be seen that children with private health insurance have the highest test scores and that those with Medicaid have the lowest.

The bottom panel of Table 1.2 shows that children with different types of coverage also have different observable characteristics. On average, those in the “always private health insurance” category live in higher income households, have better home environments (measured by the Home S-F score) and have mothers that are more likely to be married, employed, and more educated than those who did not have health insurance. On average, children without health insurance live in wealthier households than children with Medicaid coverage but in poorer households than children with private health insurance. This suggests that children with no health insurance do not meet

⁸ Sample means for children who had a given type of coverage for some periods, but not the others are not presented to save space, but are available upon request.

Medicaid eligibility requirements, but at the same time do not live in households that can afford to buy private insurance coverage.

In order to implement family fixed effects as an estimation strategy, it is important to have within family variation in health insurance coverage. Table 1.3 shows the number of families in which children have differences in the amount of time they had health insurance coverage of a given type during early childhood. There are 326 families (out of 1,461 in the sample) where children have sibling differences in the amount of time they had health insurance coverage during early childhood. The number of families in which there is within family variation in private health insurance coverage and Medicaid coverage is 444 and 331, respectively.

Results from Multivariate Models

Health Insurance and Children's Cognitive Outcomes

Table 1.4 reports estimates from multivariate regression models of the impact of average health insurance coverage (with no distinction between public or private insurance) during early childhood on each of the four cognitive measures previously described. All specifications include the controls listed at the bottom of the Table. Odd numbered columns show ordinary least squares results and even numbered columns show family fixed effects estimates. As mentioned before, the sample only includes children who have one or more siblings also interviewed by the survey during early and late childhood. Observations in the regression analyses are not weighted.

OLS results show that children who always have health insurance coverage during early childhood have test scores in math (PIAT-M), reading recognition (PIAT-R), reading comprehension (PIAT-C) and vocabulary (PPVT) that are 0.126, 0.127, 0.102, and 0.099 standard deviations (SD) higher than uninsured children, respectively. However, only the result for PIAT-M is statistically distinguishable from zero. Note that family fixed effects coefficient estimates are larger than OLS indicating that OLS estimates are downwards biased due to unobserved heterogeneity. This is confirmed by F-tests that reject the null of no family fixed effects and Hausman tests that reject the null of no correlation between the error term and the regressors. Family fixed effects estimates show that children who always had health insurance coverage during early childhood have PIAT-M, PIAT-R, PIAT-C, and PPVT test scores that are respectively 0.229, 0.146, 0.236, and 0.169 SDs higher than their uninsured siblings.⁹ All of these results are statistically significant. Dividing these coefficient estimates by five, the length in years of the early childhood period, gives us an estimate of the effect of an additional year of coverage. This transformation reveals that one year of health insurance coverage during early childhood increases test scores by 0.03 to 0.05 SDs. One implicit assumption of this transformation is that every year of coverage has the same effect on test scores.

To investigate if the effect of health insurance on test scores varies depending on the type of coverage children have I replace the aggregate measure of health insurance coverage during early childhood by measures of private health insurance and Medicaid coverage. Table 1.5 shows the OLS and FE coefficient estimates for these variables.

⁹ Not reported estimates of contemporaneous health insurance coverage are also positive.

Family fixed effects estimates show that children who have private health insurance coverage every survey wave during early childhood have PIAT-M, PIAT-R, PIAT-C and PPVT scores that are 0.219, 0.160, 0.224 and 0.171 SDs higher than their uninsured siblings. These coefficient estimates are statistically different from zero, except for the estimate for PPVT scores. Coefficient estimates for Medicaid are very similar to the ones obtained for private health insurance and t-tests cannot reject the null hypothesis that both types of coverage have the same effect on test scores.

It is possible that the effect of early childhood health insurance coverage diminishes, as children get older. To investigate this hypothesis I divide the late childhood period in two sub-periods, ages 5-9 and 10-14. The top panel of Table 1.6 shows fixed effects coefficient estimates of health insurance coverage during early childhood on test scores taken at ages 5-9. Regressions include the same controls listed at the bottom of Table 1.4. Averages of time-varying characteristics are calculated over ages 0-4 and 5-9. Controls for number of interviews at those ages are also included. The results for PPVT scores are not shown because as discussed in the data section the majority of children are tested when they were 10 years old or older. The results in the top panel of Table 1.6 show that the scores of PIAT-M, PIAT-R and PIAT-C taken at ages 5-9 are 0.204, 0.195 and 0.190 SDs higher for siblings who always had health insurance coverage during early childhood. These coefficient estimates are very similar in magnitude to the ones obtained in Table 1.4 and are also statistically different from zero.

The bottom panel of Table 1.6 shows fixed effects results for health insurance coverage on test scores taken at ages 10-14. Regressions include the same controls

described at the bottom of Table 1.4 and dummy variables that control for the number of times the child was interviewed at each age range.¹⁰ The effect of health insurance coverage at ages 0-4 on PIAT-M, PIAT-R and PIAT-C taken at ages 10-14 is smaller than the effect on test scores taken at ages 5-9; however the confidence intervals in both regressions overlap, suggesting that the effect of health insurance coverage during early childhood has a lasting effect on test scores. Health insurance coverage at ages 5-9 also has a positive effect on test scores taken at ages 10-14 and t-tests cannot reject the null hypothesis that the effect of health insurance coverage at ages 0-4 is equal to the effect of health insurance coverage at ages 5-9 except for PPVT scores.¹¹ Taken together these results suggest that past period health insurance coverage has a positive effect on children's test scores and that early childhood health insurance coverage has a persistent effect on children's cognitive outcomes.

Robustness Checks

As mentioned before, even though all the econometric specifications include extensive controls for time-varying characteristics of the family, it is still possible that there are unobservable characteristics that could bias the coefficient estimates. In unreported regressions I test the sensitivity of the results reported in Table 1.4 to an alternative specification that includes interactions between the variables that measure the

¹⁰ Among children who were interviewed at least once when they were 5-9 years old 84 percent was interviewed twice and 8 percent was interviewed three times. Among children who were interviewed at least once when they were 10-14 years old 37 percent was interviewed twice and 51 percent was interviewed three times.

¹¹ The sample sizes are smaller than the ones used to estimate the results showed in Table 1.5 because there are some children who were not interviewed at ages 5-9 or at ages 10-14. Similar results were obtained when the same sample was used to estimate the regressions described at the top and bottom panels of Table 1.6.

marital and employment status of the mother. Changes in these variables could potentially result in variations in health insurance coverage and it is possible that there are interactions between these variables that affect both health insurance status and test scores. The results obtained using this alternative specification are very similar to the ones described in Table 1.4.¹²

Heterogeneity

Table 1.7 shows the effect of health insurance coverage during early childhood separately for girls and boys. The samples consist of families that have two or more children of a given gender. The medical literature has documented that morbidity and mortality rates in early life are higher for boys than for girls (Wells, 1999; Copper et al. 1994; Read et al. 1997); thus, if boys are more likely to need medical attention to treat or prevent health conditions they may also benefit more from having health insurance coverage. Results on Table 1.7 seem to confirm this hypothesis; health insurance during early childhood has a larger effect on boys' PIAT-M, PIAT-R and PIAT-C scores. Interestingly, early childhood health insurance has a larger effect on PPVT scores for girls.

¹² Another concern is that parents obtain health insurance coverage for their children in response to the child-specific propensity to get sick. As mentioned before, this is a bigger concern for the Medicaid coefficient estimates. To investigate if this is a serious concern for children in the sample I use the information on children's age, family income, household size, year of interview and state of residence along with the Medicaid eligibility rules provided by Hoynes and Luttmer (2011) to impute children's Medicaid eligibility status for every year that children were surveyed by the CNLSY. Using this information I tried to instrument children's Medicaid enrollment during early and late childhood with their Medicaid eligibility during these periods. An instrumental variables estimator will provide unbiased estimates of Medicaid enrollment if Medicaid eligibility is exogenous from children's underlying health. Unfortunately, Medicaid eligibility is a weak instrument for Medicaid enrollment once observable characteristics and family fixed effects are accounted for.

Table 1.8 shows FE results when the sample is divided according to average family income, expressed as a percentage of the Federal Poverty Line (FPL), during early childhood. The top panel uses 100 percent of the FPL as the cutoff point, the middle panel uses 200 percent, and the bottom panel uses 300 percent as the cutoff point. Families in the sample are those who had two or more children ages 0-4 when their family income was below (or above) a given threshold. For children younger than 5, the average income threshold to be eligible for Medicaid between 1988 and 2008 was 170 percent of the FPL. The middle panel of Table 1.8 shows that among children living in households with incomes less or equal to 200 percent of the FPL the effect of health insurance coverage ranges between 0.074SD-0.249SD. These estimates are larger than the ones for children living in households with incomes greater than 200 percent of the FPL, except for PPVT. The use of alternative cutoff points show that for children in low-income households the effect of health insurance coverage is larger on PIAT-M and PIAT-R scores than for children in the rest of the sample. This is not always the case for PIAT-C and PPVT scores.

Discussion

The evidence presented in this paper suggests that health insurance coverage during early childhood has a lasting effect on children's cognitive outcomes. Both private health insurance and Medicaid coverage at ages 0-4 have positive and statistically significant effects on test scores taken at ages 5-14. Previous investigations have found that health insurance coverage improves children's health and that children who are in better health have better cognitive and educational outcomes. This is the first paper that

uses individual level data to find a reduced form relationship between health insurance coverage and children's cognitive outcomes. Levine and Schazzenback (2009) found a positive relationship between Medicaid expansions at birth and state-average reading test scores of fourth-grade children. This investigation expands their analysis by using longitudinal, individual level data, information on private health insurance and Medicaid coverage, and a wider range of measures of cognitive outcomes.

The strengths of this study are that the CNLSY data allows controlling for unobserved family heterogeneity that may be correlated with health insurance coverage and with children's test scores. Also information on the CNLSY permits me to control for a detailed list of time-varying covariates. Fixed effects models could be biased if there is remaining unobserved family or child heterogeneity. However, sensitivity tests show that the results are robust to additional measures for family characteristics.

There are several limitations that should be recognized. First, there are several sample restrictions that had to be imposed to the CNLSY in order to analyze children's long-term outcomes and to be able to implement the fixed effects methodology. Children in the estimating sample live in households that have higher income, mothers in the estimated sample are older, more educated, and more likely to be married, children are more likely to have health insurance coverage in general, and more likely to have private health insurance in particular. The evidence presented in this paper suggests that the benefits of health insurance coverage are larger for children living in low-income households; thus, we could expect that the effects of health insurance coverage will be larger for the complete CNLSY sample. Second, the length of the survey and the size of

the sample limited the type of outcomes that could be analyzed. It would be interesting to analyze the effect of health insurance in other school related outcomes, such as grade repetition or high school graduation, and also examine longer term outcomes such as earnings and employment status. However, previous evidence shows that the cognitive outcomes analyzed in this paper (PIATs) are correlated with educational attainment, employment, and wages (Heckman et al., 2006). This would suggest that the positive effect of health insurance on PIATs scores might also impact future educational and labor market outcomes. Finally, the availability of more detailed data on health outcomes would allow a better understanding of the mechanisms behind the relationship between health insurance and cognitive outcomes.

Tables and Figures

Table 1.1 - Distribution of Children According to the Proportion of Time they were Covered by Health Insurance.

	Health insurance coverage (any type)	Private health insurance coverage	Medicaid coverage
Early childhood (ages 0-4)			
Always	84.6	65.9	9.7
Sometimes	12.5	17.7	14.2
Never	2.9	16.4	76.1
Late childhood (ages 5-9)			
Always	86.6	68.7	8.2
Sometimes	11.6	18.4	13.8
Never	1.8	12.9	78.0
Late childhood (ages 10-14)			
Always	90.1	74.9	7.3
Sometimes	7.1	13.2	9.9
Never	2.8	11.9	82.8

The table shows information for CNLSY siblings that were interviewed during early and late childhood. The categories “always” mean that every time the child was surveyed by the CNLSY the mother reported that child had that specific type of health insurance. Similarly, the category “never” means that every time the child was surveyed by the CNLSY the mother reported the child was not covered by that specific type of health insurance. The categories “sometimes private health insurance” and “sometimes Medicaid” overlap because children could transition from having private health insurance to Medicaid coverage or vice versa.

Table 1.2 – Means of Outcomes and Observable Characteristics According to Health Insurance Coverage during Early Childhood (Ages 0-4). Sample of Siblings.

	All children	Always health insurance (any type)	Always private health insurance	Always Medicaid	Always uninsured
Outcomes					
Average scores of test taken at ages 5-14					
PIAT-M (Mathematics)	10.02	10.08***	10.25***	9.36***	9.70
PIAT-R (Reading Recognition)	10.01	10.07**	10.23***	9.39***	9.71
PIAT-C (Comprehension)	9.97	10.03*	10.20***	9.34***	9.74
PPVT (Vocabulary)	10.09	10.16***	10.35***	9.37***	9.75
Time-Varying Controls (early childhood)					
Household					
Income (in 10,000\$)	6.96	7.58***	9.18***	1.50***	3.36
Home Simple Form score	9.99	10.04	10.27***	9.14***	9.84
Number of people ages 0-5	0.74	0.72***	0.67***	0.94	0.86
Number of people 6-11	0.61	0.60	0.50	0.93***	0.60
Number of people 12-17	0.21	0.20	0.14	0.38***	0.17
Number of adults	2.01	1.99**	2.04	1.75***	2.13
Mother					
AFDC participation	0.16	0.15	0.02***	0.72***	0.15
Married	0.75	0.76	0.90**	0.23***	0.79
Single	0.14	0.14	0.04*	0.48***	0.10
Divorced or widow	0.12	0.10	0.06**	0.29***	0.11
Married with employed spouse	0.68	0.70	0.85***	0.12***	0.72
Employed	0.57	0.58**	0.67***	0.16***	0.54
High school drop out	0.17	0.14**	0.06***	0.54***	0.34
High school graduate	0.35	0.33**	0.32***	0.33**	0.44
Has college education	0.48	0.52***	0.62***	0.13*	0.21
Time-Varying Controls (late childhood)					
Household					
Income (in 10,000\$)	7.43	8.14***	9.86***	1.45***	3.35
Home Simple Form score	10.01	10.07**	10.34***	8.82***	9.71
Number of people ages 0-5	0.53	0.52	0.47	0.66**	0.42
Number of people 6-11	0.75	0.74	0.70	0.76	0.72
Number of people 12-17	0.57	0.57	0.51*	0.88*	0.66
Number of adults	1.99	1.99	2.04	1.76***	2.15
Mother					
AFDC participation	0.12	0.10	0.02*	0.49***	0.10

Married	0.69	0.71	0.84***	0.18***	0.65
Single	0.10	0.10	0.03*	0.38***	0.12
Divorced or widow	0.20	0.19	0.13	0.44***	0.23
Married with employed spouse	0.62	0.65**	0.78***	0.08***	0.51
Employed	0.68	0.70	0.76**	0.37***	0.61
High school drop out	0.16	0.14	0.05**	0.54***	0.21
High school graduate	0.33	0.31**	0.30	0.33	0.34
Has college education	0.51	0.55	0.65**	0.13***	0.45

Time Invariant Controls

Mother

Age at time of child's birth	28.9	29.3***	29.6***	26.1	26.1
Cigarettes during pregnancy	0.27	0.25	0.19	0.52**	0.32
Drinks during pregnancy	0.56	0.58*	0.59	0.66*	0.42

Child

Female	0.49	0.49	0.50	0.46	0.52
Child is non-white	0.46	0.43***	0.33***	0.85***	0.67
Fetal growth (a)	2.56	3.07**	3.11	2.91***	3.17
Birth order	2.36	2.33	2.10	3.10***	2.27
Age at last interview	12.5	12.4**	12.4**	12.6	12.8

Asterisks indicate whether the means are significantly different for children with health insurance coverage to those without coverage. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. (a) Fetal growth=birth weight/length of pregnancy. Test scores were normalized to have a mean of 10 and a standard deviation of 1. Information on observables characteristics during late childhood is not presented to save space. This information is available upon request.

Table 1.3 – Number of Families and Children with Variation in Health Insurance Coverage at Ages 0-4.

	No. of children with differences in coverage	No. of families with differences in coverage
Health insurance (overall)	838	326
Private health insurance	1,157	444
Medicaid	892	331

The total sample size is 3,508 children, who belong to 1,461 families.

Table 1.4 – The Effect of Early Childhood Health Insurance Coverage on Children’s Test Scores Taken at Ages 5-14. Sample of Siblings. OLS and Family Fixed Effects Results.

	PIAT-M		PIAT-R		PIAT-C		PPVT	
	OLS	Fixed Effects	OLS	Fixed Effects	OLS	Fixed Effects	OLS	Fixed Effects
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Avg. Health insurance coverage, ages 0-4	0.126*	0.229***	0.127	0.146*	0.102	0.236***	0.099	0.169*
	(0.071)	(0.086)	(0.078)	(0.087)	(0.071)	(0.083)	(0.077)	(0.102)
Observations	3,507	3,507	3,508	3,508	3,372	3,372	2,979	2,979
Number of families		1,460		1,461		1,410		1,253

*** p<0.01, ** p<0.05, * p<0.1. Clustered standard errors in parentheses. Every column includes controls for children’s sex, race, dummies for birth order, dummies for birth of year, dummies for the number of times the child was interviewed during early and late childhood, fetal growth, number of cigarettes smoked during pregnancy, number of alcoholic drinks consumed during pregnancy, mother’s age at the time of the child’s birth. Time-varying controls include: average real household income (linear, quadratic and cubic), mother’s average: education, employment and marital status, average employment status of the spouse if the mother is married, average welfare participation, average household demographics (i.e. number of people ages 0-5, 6-11, 12-17, and 18 and older), average test scores obtained in the Home Observation Measurement of the Environment Short Form (HOME-SF), average child’s age in months (linear and quadratic). Time-varying controls are averages calculated over the early (ages 0-4) and late (ages 5-14) childhood periods. Regressions also include controls for average health insurance coverage during late childhood. The dependent variables are average scores of test taken at ages 5-14. Test scores are normed to have a mean of 10 and a standard deviation of 1.

Table 1.5 – The Effect of Private Health Insurance and Medicaid Coverage on Children’s Test Scores Taken at Ages 5-14. Sample of Siblings.

	PIAT-M		PIAT-R		PIAT-C		PPVT	
	OLS	Fixed Effects	OLS	Fixed Effects	OLS	Fixed Effects	OLS	Fixed Effects
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Avg. Private Insurance coverage, ages 0 - 4	0.100 (0.074)	0.219** (0.091)	0.118 (0.081)	0.160* (0.093)	0.067 (0.075)	0.224** (0.087)	0.095 (0.079)	0.171 (0.112)
Avg. Medicaid coverage, ages 0 - 4	0.165* (0.087)	0.218** (0.111)	0.153 (0.098)	0.113 (0.105)	0.149* (0.085)	0.218** (0.109)	0.043 (0.101)	0.093 (0.125)
Observations	3,507	3,507	3,508	3,508	3,372	3,372	2,979	2,979
Number of families		1,460		1,461		1,410		1,253

*** p<0.01, ** p<0.05, * p<0.1. Clustered standard errors in parentheses. Every column includes the same controls described at the bottom of Table 1.4. The dependent variables are average scores of test taken at ages 5-14. Test scores are normed to have a mean of 10 and a standard deviation of 1.

Table 1.6 – The Effect of Health Insurance Coverage on Children’s Test Scores. Family Fixed Effects Results.

	PIAT-M (1)	PIAT-R (2)	PIAT-C (3)	PPVT (4)
Test scores taken at ages 5-9				
Avg. Health Insurance coverage (any type), ages 0-4	0.204** (0.095)	0.195** (0.091)	0.190* (0.105)	
Observations	3,387	3,388	3,079	
Number of families	1,412	1,413	1,296	
Test scores taken at ages 10-14				
Avg. Health Insurance coverage (any type), ages 0-4	0.209 (0.127)	0.087 (0.115)	0.129 (0.102)	0.021 (0.109)
Avg. Health Insurance coverage (any type), ages 5-9	0.173 (0.142)	0.235* (0.141)	0.219* (0.126)	0.323** (0.129)
Observations	2,753	2,752	2,723	2,253
Number of families	1,170	1,169	1,159	967

*** p<0.01, ** p<0.05, * p<0.1. Clustered standard errors in parentheses. Every column includes the same controls described at the bottom of Table 1.4. Regressions in the top panel control for the number of times children were interviewed at ages 0-4, 5-9, private health insurance and Medicaid coverage at ages 5-9. Time-varying controls are averaged over ages 0-4 and ages 5-9. The dependent variables in the top panel of the table are average scores of test taken at ages 5-9. Regressions in the bottom panel control for the number of times children were interviewed at ages 0-4, 5-9, 10-14, private health insurance and Medicaid coverage at ages 10-14. Time-varying controls are averaged over early (ages 0-4) and late (ages 5-14) childhood. The dependent variables in the bottom panel of the table are average scores of test taken at ages 10-14. Test scores are normed to have a mean of 10 and a standard deviation of 1.

Table 1.7 - The Effect of Health Insurance on Children's Test Scores Taken at Ages 5-14. Girls Versus Boys. Family Fixed Effects Results.

	PIAT-M		PIAT-R		PIAT-C		PPVT	
	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys
Avg. Health Insurance coverage, ages 0-4	0.014	0.481***	-0.041	0.409**	0.074	0.481**	0.289*	0.163
	(0.176)	(0.168)	(0.161)	(0.191)	(0.149)	(0.190)	(0.159)	(0.251)
Observations	973	1,106	973	1,104	925	1,048	821	931
Number of families	438	501	438	500	415	476	371	425

*** p<0.01, ** p<0.05, * p<0.1. Clustered standard errors in parentheses. Every column includes the same controls described at the bottom of Table 4. The dependent variables are average scores of test taken at ages 5-14. Test scores are normed to have a mean of 10 and a standard deviation of 1.

Table 1.8 - The Effect of Health Insurance on Children's Test Scores Taken at Ages 5-14. Sample Divided According to Average Family Income during Early Childhood. Family Fixed Effects Results.

	PIAT-M		PIAT-R		PIAT-C		PPVT	
	Average family income during early childhood							
	<=100% fpl	>100% fpl	<=100% fpl	>100% fpl	<=100% fpl	>100% fpl	<=100% fpl	>100% fpl
Avg. Health Insurance coverage, ages 0-4	0.232	0.191*	0.232	0.191*	0.111	0.124	0.070	0.192*
	(0.200)	(0.101)	(0.200)	(0.101)	(0.205)	(0.122)	(0.217)	(0.114)
Observations	720	2,489	721	2,489	683	2,400	651	2,080
Number of families	280	1,075	281	1,075	268	1,038	259	905
	<=200% fpl	>200% fpl	<=200% fpl	>200% fpl	<=200% fpl	>200% fpl	<=200% fpl	>200% fpl
Avg. Health Insurance coverage, ages 0-4	0.223**	0.150	0.196*	0.020	0.249**	0.071	0.074	0.122
	(0.113)	(0.175)	(0.114)	(0.194)	(0.115)	(0.209)	(0.150)	(0.183)
Observations	1,341	1,752	1,342	1,752	1,280	1,686	1,171	1,450
Number of families	532	767	533	767	510	740	474	638
	<=300% fpl	>300% fpl	<=300% fpl	>300% fpl	<=300% fpl	>300% fpl	<=300% fpl	>300% fpl
Avg. Health Insurance coverage, ages 0-4	0.268***	0.063	0.165*	0.115	0.238**	0.630*	0.161	0.095
	(0.099)	(0.276)	(0.098)	(0.360)	(0.094)	(0.341)	(0.118)	(0.355)
Observations	2,041	1,108	2,042	1,108	1,959	1,063	1,776	893
Number of families	837	495	838	495	807	477	738	402

*** p<0.01, ** p<0.05, * p<0.1. Clustered standard errors in parentheses. Every column includes the same controls described at the bottom of Table 1.4. The dependent variables are average scores of test taken at ages 5-14. Test scores are normed to have a mean of 10 and a standard deviation of 1.

CHAPTER II

DOES UNIVERSAL COVERAGE IMPROVE HEALTH? THE MASSACHUSETTS
EXPERIENCE

(Co-authored with Charles Courtemanche)

Abstract

In 2006, Massachusetts passed health care reform legislation designed to achieve nearly universal coverage through a combination of insurance market reforms, mandates, and subsidies that later served as the model for national health care reform. We provide evidence that the Massachusetts reform improved self-assessed overall health, physical health, mental health, functional limitations, joint disorders, body mass index, and moderate physical activity. The effect on overall health was strongest among women, minorities, near-elderly adults, and those with low incomes. Using the reform to instrument for health insurance coverage, we estimate a sizeable positive impact of coverage on health.

Introduction

A major objective of the Patient Protection and Affordable Care Act (ACA) signed into law in March of 2010 is to increase health insurance coverage in the United States to nearly universal levels through a combination of insurance market reforms, mandates, and subsidies. Although the law survived constitutional challenges, it remains at the center of political debate, with possibilities remaining for full or partial repeal or denial of financing during the budgetary process. This ongoing debate highlights the need for projections of the law's impacts on health, health care utilization, and state and federal budgets. The multi-faceted nature of the reform and breadth of the population affected suggests that evidence from coverage expansions in other contexts, such as Medicaid, will be of only limited usefulness.

The most similar intervention to date to the ACA is the Massachusetts health care reform of April 2006, entitled "An Act Providing Access to Affordable, Quality, Accountable Health Care" and commonly called "Chapter 58" (Long, 2008).¹³ The law enabled Massachusetts to lower its uninsurance rate to 2% by 2010 through a strategy called "incremental universalism," or "filling the gaps in the existing system ... rather than ripping up the system and starting over" (Massachusetts' Division of Health Care, Finance and Policy, 2010; Gruber, 2008a:52). Gruber (2010) describes Massachusetts' approach to incremental universalism as involving a "three legged stool" of insurance market reforms, mandates, and subsidies (Gruber, 2010).

¹³ For a more detailed description of the law, see Long (2008), McDonough et al. (2006) and Gruber (2008a, 2008b).

The first leg of the stool reforms non-group insurance markets in an effort to ensure the availability of coverage for those without access to employer-provided or public insurance. Insurers are not allowed to deny or drop coverage based on pre-existing conditions (guaranteed issue) or vary premiums to reflect health status aside for limited adjustments for age and smoking status (community rating) (Kirk, 2000; McDonough et al., 2006). A health insurance exchange, the Commonwealth Health Insurance Connector Authority, offers plans developed by licensed health insurance companies for those without access to group markets. Enrollment on the Connector began in October 2006 for those with incomes below 100% of the federal poverty line (FPL), in January 2007 for those up to 300% FPL, and in May 2007 for everyone else. Additionally, private health insurance plans are required to provide coverage for young adults on their parents' plans for up to two years after they are no longer dependents or until their 26th birthday (McDonough et al., 2006).¹⁴

This first leg alone would likely lead to adverse selection and a “death spiral” with rising premiums gradually driving healthy individuals out of the non-group market. The second leg of the three-legged stool therefore involves mandates requiring adults to be covered by health insurance and employers to provide health insurance. Individuals without adequate coverage face a penalty of half of the lowest premium they would have paid in a Health Connector-certified plan. Employers with more than 10 employees must make a “fair and reasonable” contribution toward an employer health insurance plan or

¹⁴ Guaranteed issue and community rating have been in place in Massachusetts since 1996. The 1996 law only allowed premiums to vary with age and geography; Chapter 58 further allowed them to vary with tobacco use. The insurance exchange and the requirement regarding young adults on their parents' plans both started with Chapter 58.

pay a state assessment of up to \$295 per full-time equivalent worker per year (Massachusetts Health Insurance Connector Authority, 2008).¹⁵ The mandates took effect in July 2007.

To help low- and middle-income households be financially able to comply with the mandate, the third leg of the Massachusetts reform provides subsidies and Medicaid expansions. Chapter 58 specifies that health insurance be free for people below 150% FPL and that premiums be subsidized on a sliding scale for those between 150% and 300% FPL with no deductibles.¹⁶ The reform also expands Medicaid to cover children below 300% FPL (McDonough et al., 2006).

Taking into account the costs of the subsidies and Medicaid expansions as well as the savings from reduced safety net payments, Raymond (2009) estimates the annual fiscal cost of the reform to be \$707 million. Through a waiver allowing for a more flexible use of federal Medicaid matching money, half of this amount comes from the federal government, leaving the state government's share at \$353 million.

Table 2.1 compares Massachusetts' approach to incremental universalism with that of the Affordable Care Act.¹⁷ Though there are differences in some of the details, both the Massachusetts and national reforms were clearly motivated by the same "three-

¹⁵ Minimum requirements plans must meet to satisfy the mandates include coverage for prescription drugs and preventive and primary care, as well as maximums on deductibles and out-of-pocket spending.

¹⁶ For instance, in 2008 a family with an income between 150% and 200% of the poverty line paid a premium of \$35 per adult, while a family with an income in the 250% to 300% range paid \$105 per adult.

¹⁷ Coverage expansion was the primary focus of both the Massachusetts and national reforms. However, the national reform was more comprehensive, consisting of nine titles that each had their own reform agenda: I. Insurance Coverage, II. Medicaid and the Children's Health Insurance Program, III. Delivery System Reform, IV. Prevention and Wellness, V. Workforce initiatives, VI. Fraud, Abuse and Program Integrity, VII. Biologic Similar, VIII. Community Living Assistance Services and Supports, IX. Revenue Provisions (Patel and McDonough, 2010).

legged stool” approach to incremental universalism. Both featured guaranteed issue, community rating, insurance exchanges, mandates, Medicaid expansions, and subsidies. For these reasons, analyzing the effects of health care reform in Massachusetts provides the best available predictor to date of the implications of the Affordable Care Act.

Given that recent nature of the Massachusetts reform, researchers are only beginning to understand its impacts. Long et al. (2009) find that by 2008 the uninsured rate decreased by 6.6 percentage points for the overall nonelderly population and 17.3 percentage points for lower-income adults.¹⁸ Long and Stockely (2011) find a decrease in unmet medical needs because of cost among lower income adults but also some evidence of delays in care from being unable to find a provider. Yelowitz and Cannon (2010) show that Chapter 58’s impact on coverage was mitigated by the crowding out of private insurance. They also investigate the reform’s effect on self-assessed health, finding mixed results: an increase in the probability of reporting at least good health but a decrease in the probability of reporting at least very good health. Cogan et al. (2010) estimate that the reform increased employer-sponsored insurance premiums by about 6%. Kolstad and Kowalski (2010) show that the reform reduced levels of uninsurance by 36% among the population of hospital discharges. Length of stay and the number of inpatient admissions originating from the emergency room both decreased, with some evidence also suggesting an increase in the utilization of preventive services, a decline in hospitalizations for preventable conditions, and an improvement in quality of care. Miller (2011a) finds a reduction in non-urgent emergency room visits, consistent with the

¹⁸ These results support preliminary evidence found by Long (2008) using information from 2006 and 2007.

newly-insured having access to such care in other settings. Miller (2011b) focuses on children's outcomes, finding a substitution from emergency room care to office visits, a reduction in medical needs unmet because of cost, and an increase in the probability of reporting excellent health. Kowalski and Kolstad (2012) exploit the reform's effect on employer-provided health insurance to show that wage reductions almost completely offset the cost of health insurance benefits.

We contribute to this growing literature by examining Chapter 58's effect on the self-assessed health of adults. Though many open questions remain about the reform's effectiveness, as Gruber (2011b:190) writes, "the most significant of these is the impact of reform on the health of citizens." We utilize individual-level data from the Behavioral Risk Factor Surveillance System (BRFSS), which allows for the use of longer pre- and post-treatment periods, a much larger sample, and a broader range of health-related questions than Yelowitz and Cannon (2010), enabling us to obtain clearer results.¹⁹

First, an ordered probit difference-in-differences analysis shows that the reform increased the probability of individuals reporting excellent or very good health while reducing their probability of reporting good, fair, or poor health. A variety of robustness checks and placebo tests support a causal interpretation of the results. The estimates suggest that annual government spending for each adult transitioned into excellent or very good health is \$9,827, split evenly between the Massachusetts and federal

¹⁹ Specifically, Yelowitz and Cannon (2010) use Current Population Survey supplements and compare a pre-treatment period of 2005-2006 with a post-treatment period of 2008. They conduct a difference-in-differences analysis with other New England states as controls. Their sample size is 41,873. In contrast, we utilize data from 2001-2010 and have a sample size of 2,879,296 in our main analysis and 340,592 when we restrict the sample to New England.

governments. We then provide evidence that the reform improved a number of determinants of overall self-assessed health: physical health, mental health, functional limitations, joint disorders, body mass index, and moderate physical activity. Next, we examine heterogeneity and find that the reform's effect on overall health was strongest for women, minorities, near-elderly adults, and those with incomes low enough to qualify for the law's subsidies. Notably, the estimates imply a 24% reduction in the disparity in self-reported health between blacks and whites. Finally, we exploit the plausibly exogenous variation in coverage created by the reform to estimate that obtaining health insurance leads to a large improvement in health.

Health Insurance and Health

An important part of the argument for universal coverage is the assumption that health insurance improves health. As quoted by Yelowitz and Cannon (2010), Levy and Meltzer (2008) write,

The central question of how health insurance affects health, for whom it matters, and how much, remains largely unanswered at the level of detail needed to inform policy decisions. ... Understanding the magnitude of health benefits associated with insurance is not just an academic exercise ..., it is crucial to ensuring that the benefits of a given amount of public spending on health are maximized (p. 400).

This section provides a brief summary of theoretical and empirical research on the topic and summarizes our contribution to this broader literature.

Grossman (1972) models health as a durable capital stock that is also an input in the production of healthy time. Health capital depends on the initial endowment of health, past period health, and past period investments made to preserve it. Medical care and

time spent in health producing activities are the main forms of health investment. Every period people face uncertainty as to whether they will be affected by a negative health shock, so they buy health insurance to protect themselves against unexpected medical costs. Because health insurance reduces the price of care faced by the consumer it increases the demand for medical care (Arrow, 1963; Pauly, 1968). This increase in consumption of care could result in better health, but if the additional medical care is redundant health outcomes may remain the same or even deteriorate. This effect is sometimes known as “flat of the curve” medical care, because diminishing returns in the health production function imply that at some point the health gains associated with more medical care may be very small (Doyle, 2005).

The majority of empirical investigations into the relationship between health insurance and health are observational studies that use multivariate regression analysis. A review of these studies by Hadley (2003) shows that 15 out of the 20 published between 1991 and 2001 found a positive association between health insurance coverage and recovery from health conditions such as cancer, trauma, and appendicitis. Health insurance was also associated with better overall health status and lower mortality risk in all of the studies that examined these outcomes. However, these relationships cannot be interpreted as causal because the research designs did not address the potential for unobserved heterogeneity and reverse causality.

During the 1970’s the RAND Health insurance experiment randomly assigned families to health insurance plans with coinsurance rates ranging from 0% to 95%, with all medical expenses covered over a threshold. Medical care use increased among people

assigned to plans with lower coinsurance rates, but health outcomes only improved among the poor (Manning et al., 1987). However, this experiment only shows the impact of health insurance along the intensive margin from less to more generous coverage, not the extensive margin of no coverage to any coverage. It is also unclear to what extent findings from the 1970s are applicable today.

Some studies have taken advantage of the plausibly exogenous variation provided by public insurance programs like Medicaid and Medicare in order to address the endogeneity of coverage. Currie and Gruber (1996a, 1996b) find that Medicaid expansions decrease infant mortality and low birth weight, while Dafny and Gruber (2005) show that they also reduce avoidable hospitalizations among children. Most recently, Finkelstein et al. (2011) exploit a 2008 Oregon lottery in which winners were given the chance to apply for Medicaid to show that coverage improves self-reported physical and mental health. The randomization allows for clean identification of the causal effects of Medicaid eligibility, at least among the low-income uninsured lottery participants.

Evidence on the effect of Medicare on the health of seniors is mixed. Card et al. (2004) find that obtaining Medicare coverage at age 65 improves the self-assessed health of Hispanics and people with low levels of education; however, the effect for the whole sample is smaller and insignificant. Finkelstein and McKnight (2008) show that 10 years after the introduction of Medicare there was not a statistically significant impact on mortality rates for people older than 65. Card et al. (2009) find more favorable results: a

reduction in the 7-day mortality rate among emergency room patients older than 65 compared to those right below that cutoff.

A few studies attempt to estimate the causal effect of insurance on health in contexts other than public programs, again finding mixed results. Pauly (2005) uses marital status and firm size as instruments for private insurance coverage and finds a positive but insignificant effect of insurance on self-reported health and a negative but insignificant effect on the probability of having a chronic condition. Doyle (2005) shows that uninsured patients receive less medical care and have higher mortality rates than insured patients after a random health shock (a car accident).

To summarize, the extant literature suggests that health insurance coverage appears to improve health in some contexts but not others. The uninsured in the U.S. consist of a number of groups, including those too sick to obtain coverage, those too healthy to feel insurance is necessary, and those too poor to afford private coverage but not poor enough to qualify for public insurance programs. Any attempt at universal coverage in the U.S. will therefore involve coverage expansions across a highly heterogeneous group, making it unclear the extent to which these prior findings are applicable. The Massachusetts health care reform provides a unique opportunity to examine an intervention that affects a large portion of the uninsured population.

Data

Health summarizes a combination of factors that reflect physical and mental well-being. Among the usual indicators used to measure health in empirical investigations are mortality rates, hospitalization rates, and self-assessments of overall

health. Our study focuses on self-assessments. State-level mortality information is not currently available for a long enough time after the reform to construct an adequate post-treatment period. Even if more recent data were available, examining mortality rates alone would not capture incremental improvements in health resulting from, for instance, better treatment for a chronic but non-life threatening condition. Hospitalizations are not an appropriate measure of overall health in this context since, to the extent that hospitalizations are price sensitive, changes in hospitalizations after the reform might simply be a direct result of the lower price faced by the newly-insured rather than changes in health.

This paper uses data from the BRFSS, a telephone survey of health and health behaviors conducted by state health departments in collaboration with the Centers for Disease Control and Prevention. The BRFSS, which consists of repeated annual cross sections of randomly-sampled adults, is well suited for our analysis for several reasons. First, the dataset contains the necessary variables, including multiple self-reported health measures, demographic characteristics, and state, month, and year identifiers. Second, since the BRFSS spans 1984 to 2010 and included all 50 states plus the District of Columbia by 1995, the data cover a long enough time period to examine both post-reform outcomes and pre-reform trends. Third, the BRFSS contains an unusually large number of observations – over 2.8 million in our analysis sample of 2001 through 2010. A large sample is critical to obtaining meaningful precision when examining the impact of a state-level program with effects that might be concentrated amongst only a fraction of the population.

Our main dependent variable is a self-reported health index asking respondents to rate their overall health as poor (0), fair (1), good (2), very good (3), or excellent (4). This index has been previously used by other studies analyzing the impact of health insurance on health (Card et al., 2004; Pauly, 2005; Yelowitz and Cannon, 2010) and has been repeatedly shown to be correlated with objective measures of health such as mortality (e.g. Idler and Benyamini, 1997; DeSalvo et al., 2006; Phillips et al., 2010). According to Idler and Benyamini, another advantage of the index is that it is a global measure of health that captures the full range of diseases and limitations a person may have.

The primary concern with the self-reported health index is its subjective nature. We will be able to flexibly control for the sources of reporting heterogeneity identified in the literature, such as age, income, and gender (Ziebarth, 2010). Nonetheless, the estimated effect of the reform on self-assessed health could still reflect factors beyond objective health. For instance, improved access to medical care might increase awareness about medical conditions, causing one to self-report a lower health status after obtaining insurance coverage, *ceteris paribus* (Strauss and Thomas, 2007). In this case, the reform's effect on self-assessed health would be smaller than its effect on objective health. Alternatively, if the peace of mind from having health insurance influences one's answers to subjective health-related questions, the reform could lead to larger improvements in self-assessed health than objective health.

Consequently, we also utilize a number of other health-related dependent variables in an attempt to verify that the results for the overall self-reported health index

are not driven merely by subjectivity. First, we consider number of days out of the past 30 not in good physical health and number of days out of the past 30 not in good mental health. These variables are somewhat less subjective than the overall health index because the respondents are specifically asked to consider a particular component of health. Even less subjective is the next health measure: number of days out of the past 30 with health-related functional limitations. Our last five health-related dependent variables – an indicator for the presence of activity-limiting joint pain, body mass index (BMI), minutes per week of moderate physical activity, minutes per week of vigorous physical activity, and an indicator for whether the individual currently smokes – are quite specific and therefore the least open to subjective interpretation.^{20 21}

We measure coverage with a binary variable reflecting whether or not the individual has “any kind of health care coverage, including health insurance, prepaid plans such as HMOs, or government plans such as Medicare.” The BRFSS does not indicate the source of coverage or provide any information on premiums, deductibles, or copayments. Finally, we utilize as control variables the BRFSS’ information on age, marital status, race, income, education, marital status, and current pregnancy status.

²⁰ BMI=weight in kilograms divided by height in squared meters. Self-reported weight and height are potentially susceptible to biases. Some researchers utilize an adjustment developed by Cawley (2004) that predicts actual height and weight based on self-reported height and weight using the National Health and Nutrition Examination Survey, and then applies the prediction equation to other datasets that only include the self-reported measures. However, studies with BMI as the dependent variable have repeatedly found that applying this adjustment has little influence on the results, so we do not use it here (e.g. Courtemanche et al., 2011).

²¹ The BRFSS gives respondents guidance for how to distinguish between moderate and vigorous physical activity, reducing the subjectivity of these variables. Moderate activities include “brisk walking, bicycling, vacuuming, gardening, or anything else that causes small increases in breathing or heart rate.” Vigorous activities include “running, aerobics, heavy yard work, or anything else that causes large increases in breathing or heart rate.”

We also include four state-level variables as controls in a robustness check. The first is monthly state unemployment rate, obtained from the Bureau of Labor Statistics. Next, monthly state cigarette excise tax rates come from The Tax Burden on Tobacco (Orzechowski and Walker, 2010) and are adjusted for inflation using the Consumer Price Index for all urban consumers from the Bureau of Labor Statistics. Finally, we use annual state hospital and physician data from the Census Bureau to impute monthly estimates of numbers of hospitals and physicians per 100,000 residents.²²

Our analysis uses a ten-year window surrounding the reform, 2001 to 2010. Tables 2.2 and 2.3 compare the descriptive statistics for Massachusetts and the other states in the pre-treatment period of January 2001 through March 2006. Prior to the reform, Massachusetts was already healthier than the rest of the country along most dimensions and had a higher coverage rate. Massachusetts residents averaged higher income and more education than those in other states, and were more likely to be single and white. Massachusetts also had a relatively low unemployment rate, high cigarette tax, high physician density, and low hospital density. These baseline differences illustrate the difficulty in isolating the causal impact of Massachusetts' health care reform. A naïve estimator using only a post-treatment cross section would attribute the entire difference in health between Massachusetts and other states to the reform, including the part of the difference that was already present prior to its enactment. Our empirical analysis will therefore rely on a difference-in-differences estimator that controls for pre-treatment differences in state health as well as a number of time-varying observable characteristics.

²² Monthly estimates were calculated using the formula: $X_{estimate} = X_1 + \frac{n}{12}(X_2 - X_1)$, where X_1 and X_2 are annual estimates, and n is number of months from X_1 to $X_{estimate}$.

As a precursor to the regression analysis, Figure 2.1 plots the average values of the health status index in Massachusetts and the 50 control states (the other 49 states plus Washington, DC) every year from 2001 to 2010, along with their 95% confidence intervals. The graph also shows linear pre-treatment trends for Massachusetts and the other states, computed by regressing the mean health index on year plus a constant term. Consistent with the summary statistics from Table 2.2, Massachusetts residents had better average self-assessed health than those in the control states even before the reform. Despite this difference in baseline levels, the pre-treatment trends in both Massachusetts and the other states were both downward sloping and – critically for the validity of the difference-in-differences approach – almost exactly parallel. The year-to-year fluctuations in the control states in the pre-treatment period are estimated very precisely and lie almost exactly on top of the trend line, while the year-to-year fluctuations in Massachusetts are estimated much less precisely and deviate more substantially. This underscores the importance of utilizing a sufficiently long pre-treatment period in the regression analysis. If, for instance, 2005 – a year in which health in Massachusetts appears to have been below trend – was the only pre-treatment year, a difference-in-differences estimate might capture mean reversion in addition to the causal effect.

After the reform was passed in 2006, health in the control states remained relatively stable. In contrast, health in Massachusetts improved in 2006 – as the subsidies and Medicaid expansions took effect in the early stages of the reform’s implementation –

and again in 2009.²³ To more formally investigate whether these improvements were a causal response to health care reform, we next turn to regression analysis. The regression results will broadly support the preliminary findings from Figure 2.1, although we will see that in a regression context the health gains did not appear until 2007.

Regression Analysis

Baseline Model

We estimate the impact of Massachusetts health care reform on overall self-assessed health status using an ordered probit difference-in-differences model.²⁴ Suppose the underlying relationship between the covariates and a latent variable representing health (y^*) is given by

$$y_{ist}^* = \beta_0 + \beta_1(MA_s * During_t) + \beta_2(MA_s * After_t) + \mathbf{X}_{ist}'\boldsymbol{\beta}_3 + \sigma_s + \varphi_t + \varepsilon_{ist} \quad (2.1)$$

where i , s , and t are indices for individual, state, and month/year combination (e.g. January 2001). MA_s is a dummy variable for whether the respondent lives in Massachusetts. Following Kolstad and Kowalski (2010), we define $During_t$ as a dummy variable equal to 1 from April 2006 to June 2007, the time period after the law had been passed but before all the key provisions had been implemented. $After_t$ is a dummy

²³ Figure 2.1 may help explain the mixed results found by Yelowitz and Cannon (2010). Their pre-treatment years were 2005, in which health in Massachusetts was off its long-run trend line, and 2006, in which a causal response to the early aspects of the reform was possible. Their only post-treatment year was 2008, before the second spike in the health in Massachusetts residents seen in 2009.

²⁴ Given the strong distributional assumptions made by the ordered probit model, we also considered two more flexible approaches to modeling the impact of the reform on health. The first estimates a series of four probits with the dependent variables being indicators for fair or better, good or better, very good or better, and excellent health. The second uses the same dependent variables but estimates linear probability models. The conclusions reached are the same; the results are shown in Appendices A and B.

variable equal to 1 starting in July of 2007, when the final major component of the reform – the individual mandate – took effect. \mathbf{X}'_{ist} consists of the age, marital status, race, income, education, and pregnancy variables listed in Table 2.3. σ_s and φ_t are state and month fixed effects, while ε_{ist} is the error term.

We do not observe y_{ist}^* and instead observe an ordinal health measure y_{ist} such that

$$y_{ist} = \begin{cases} 0 & \text{if } y_{ist}^* \leq \kappa_1 \\ 1 & \text{if } \kappa_1 < y_{ist}^* \leq \kappa_2 \\ 2 & \text{if } \kappa_2 < y_{ist}^* \leq \kappa_3 \\ 3 & \text{if } \kappa_3 < y_{ist}^* \leq \kappa_4 \\ 4 & \text{if } y_{ist}^* > \kappa_4 \end{cases} \quad (2.2)$$

where κ_1 through κ_4 are constants that represent the cut-off points. An ordered probit regression of y_{ist} on the covariates from (1) computes the following probabilities of being in each of the five health states:

$$\begin{aligned} Pr(y_{ist} = 0) \\ = \Phi(\lambda_1 - \beta_1(MA_s * During_t) - \beta_2(MA_s * After_t) - \mathbf{X}'_{ist}\boldsymbol{\beta}_3 - \sigma_s - \varphi_t) \end{aligned} \quad (2.3)$$

$$\begin{aligned} Pr(y_{ist} = k) \\ = \Phi(\lambda_j - \beta_1(MA_s * During_t) - \beta_2(MA_s * After_t) - \mathbf{X}'_{ist}\boldsymbol{\beta}_3 - \sigma_s - \varphi_t) \\ - \Phi(\lambda_{j-1} - \beta_1(MA_s * During_t) - \beta_2(MA_s * After_t) - \mathbf{X}'_{ist}\boldsymbol{\beta}_3 - \sigma_s \\ - \varphi_t) \forall j \in (2,3,4) \end{aligned} \quad (2.4)$$

$$\begin{aligned} Pr(y_{ist} = 4) = 1 - \Phi(\lambda_4 - \beta_1(MA_s * During_t) - \beta_2(MA_s * After_t) - \\ \mathbf{X}'_{ist}\boldsymbol{\beta}_3 - \sigma_s - \varphi_t) \end{aligned} \quad (2.5)$$

where $\lambda_j = \kappa_j - \beta_0$, the cutoff points adjusted for the constant term. The coefficient of interest is β_2 , which captures the difference between the change in Massachusetts from the “before” to the “after” period and the change in the control states from the before to the after period – in other words, the “difference in differences.”

Computing treatment effects in non-linear models has been the source of confusion in the literature. Ai and Norton (2003) showed that the cross difference in a nonlinear model is different from the marginal effect on the interaction term, and could even be the opposite sign. However, Puhani (2008) showed that the cross difference identified by Ai and Norton (2003) is not the same as the treatment effect, and that when the treatment effect is the parameter of interest it is appropriate to focus on the coefficient of the interaction term. A similar observation has been made by Terza (2012). Following Puhani (2008), our “treatment effect on the treated” is given by

$$\tau(After = 1, MA = 1) = E[Y^1 | After = 1, MA = 1, \mathbf{X}, \varphi] - E[Y^0 | After = 1, MA = 1, \mathbf{X}, \varphi] \quad (2.6)$$

where Y^1 and Y^0 are potential outcomes with and without treatment. The “average treatment effect on the treated” is the mean of this treatment effect across those individuals living in Massachusetts in the “after” period (July 2007 through December 2009).

Because of the nonlinearity of the model, the treatment effect depends on the value of the other covariates. The effects of the reform on the probabilities of being in each of the five health states among the treated are

$$\begin{aligned} \tau_{i,MA,t}(y = 0) \\ = \Phi(\lambda_1 - \beta_2 - \mathbf{X}'_{i,MA,t}\boldsymbol{\beta}_3 - \sigma_{MA} - \varphi_t) - \Phi(\lambda_1 - \mathbf{X}'_{i,MA,t}\boldsymbol{\beta}_3 - \sigma_{MA} - \varphi_t) \end{aligned} \quad (2.7)$$

$$\begin{aligned} \tau_{i,MA,t}(y = j) \\ = [\Phi(\lambda_j - \beta_2 - \mathbf{X}'_{i,MA,t}\boldsymbol{\beta}_3 - \sigma_{MA} - \varphi_t) \\ - \Phi(\lambda_{j-1} - \beta_2 - \mathbf{X}'_{i,MA,t}\boldsymbol{\beta}_3 - \sigma_{MA} - \varphi_t)] \\ - [\Phi(\lambda_j - \mathbf{X}'_{i,MA,t}\boldsymbol{\beta}_3 - \sigma_{MA} - \varphi_t) - \Phi(\lambda_{j-1} - \mathbf{X}'_{i,MA,t}\boldsymbol{\beta}_3 - \sigma_{MA} - \varphi_t)] \\ \forall j \in (2,3,4) \end{aligned} \quad (2.8)$$

$$\begin{aligned} \tau_{i,MA,t}(y = 4) \\ = 1 - \Phi(\lambda_4 - \beta_2 - \mathbf{X}'_{i,MA,t}\boldsymbol{\beta}_3 - \sigma_{MA} - \varphi_t) \\ - [1 - \Phi(\lambda_4 - \mathbf{X}'_{i,MA,t}\boldsymbol{\beta}_3 - \sigma_{MA} - \varphi_t)] \\ = \Phi(\lambda_4 - \mathbf{X}'_{i,MA,t}\boldsymbol{\beta}_3 - \sigma_{MA} - \varphi_t) - \Phi(\lambda_4 - \beta_2 - \mathbf{X}'_{i,MA,t}\boldsymbol{\beta}_3 - \sigma_{MA} - \varphi_t) \end{aligned} \quad (2.9)$$

where the state subscript s has been replaced by MA for Massachusetts, and t is restricted to the “after” period.

The key identifying assumption in the difference-in-differences model is that $MA_s * During_t$ and $MA_s * After_t$ are uncorrelated with the error term. In other words, the estimates can be interpreted as causal effects of the reform if we assume that in the absence of the reform changes over time in health would have been the same in Massachusetts and the control states, conditional on the control variables. The similarity of Massachusetts’ pre-treatment trend in health to that of the other states shown in Figure 2.1 provides preliminary support for this assumption. We therefore use all 50 other states (49 states plus the District of Columbia) as the control group in the baseline regression, and consider several alternatives in the Robustness Checks’ Section.

Our standard errors in the baseline regression are heteroskedasticity-robust and clustered by state. As shown by Bertrand et al. (2004), conventional difference-in-

differences methods can over-reject the null hypothesis because of serial correlation even when standard errors are clustered. We therefore use more stringent standards for statistical significance than usual: 0.1%, 1%, and 5% significance levels. In the Section Tests Related to Inference we will more formally investigate whether underestimated standard errors could be driving our conclusions.

The first column of Table 2.4 reports the coefficient estimates for $MA_s * During_t$ and $MA_s * After_t$ from the ordered probit regression, along with the average treatment effects on the treated in the after period.²⁵ The interaction term $MA * During$ is statistically significant at the 1% level and its effect on health is positive, suggesting that health care reform began to improve the health of Massachusetts residents even before the reform was fully implemented. This is plausible since some provisions of the reform, such as the Medicaid expansions and subsidies for those below 100% FPL, started in 2006. The interaction term $MA * After$ is significant at the 0.1% level and its coefficient estimate is more than twice as large as that for $MA * During$. Not surprisingly, the effect of the reform strengthened once it was fully implemented. This could either represent the impact of the later components, such as the mandate, or a gradual response to the earlier components. The t-statistic for $MA * After$ is 6.5, meaning that our clustered standard errors would have to be underestimated by a factor of more than three for the result to be driven by autocorrelation.

The estimated average treatment effects show that the Massachusetts health care reform decreased the probabilities of being in poor, fair and good health and increased

²⁵ Coefficient estimates for the other covariates are available upon request.

the probabilities of being in very good and excellent health. The drops in the probabilities of being in poor, fair, and good health are 0.2, 0.5, and 0.7 percentage points, respectively, while the increases in the probabilities of being in very good and excellent health are 0.2 and 1.2 percentage points.

We next conduct two back-of-the-envelope calculations to help assess the economic significance of these estimates. The first consolidates the five treatment effects into a single measure that attempts to quantify the overall increase in health. We multiply each of the treatment effects by the value of the health status index associated with the corresponding category (0 for poor, 1 for fair, 2 for good, 3 for very good, and 4 for excellent), and then divide by the sample standard deviation. This result is an overall effect on health of 0.033 standard deviations, shown in the third-to-last row of Table 2.4.²⁶ The magnitude of the impact therefore appears modest across the entire population, but perhaps large amongst the small fraction of the population who experienced a change in coverage as a result of the reform and is likely driving the results.

The second calculation combines the estimated treatment effects with the information on the reform's costs from the introduction to compute the annual fiscal cost for each adult transitioned from poor, fair, or good health to very good or excellent health. We do this first considering total government spending (federal and state), and then using only Massachusetts' share of that spending. The former provides a more relevant projection for national health care reform, while the latter is more relevant for evaluations

²⁶ This calculation should be interpreted with caution, as it relies on the strong assumption that each incremental increase in the health index represents the same improvement in health.

of the Massachusetts reform. 1.4% of the adult population transitioned into very good or excellent health. The adult population in Massachusetts was 5,138,919 in July 2010 according to the Census, so 1.4% translates to 71,945 individuals. Since the reform cost an estimated \$707 million in FY2010, total government spending is an estimated \$9,827 per year for every adult whose health improves from poor, fair, or good to very good or excellent. Since Massachusetts splits the costs evenly with the federal government, the state spends approximately \$4,914 annually per adult transitioned into very good or excellent health. These calculations are far from complete cost-effectiveness analyses, as they ignore costs to patients and private insurers as well as benefits from consumption smoothing or improvements in children's health. They do, however, provide some information about the returns to government spending while underscoring the point that financing universal coverage at the federal level is likely to be more difficult than in Massachusetts, as matching money is not available.

Robustness Checks

This section further examines the validity of the identifying assumption of common counterfactual health trends between Massachusetts and the rest of the country by considering a number of alternative control groups and adding state-level covariates. First, we use as the control group the ten states with the most similar pre-treatment average health status indices to Massachusetts ("match on pre-treatment levels"). Second, we "match on pre-treatment trends" by running regressions of average health on year plus a constant term for each state from 2001-2005 and then choosing as the comparison group the ten states with the most similar slopes to Massachusetts. Next, we use a control

group of the ten states with the most similar pre-reform health insurance coverage rates (“match on pre-treatment coverage”).²⁷ We then consider a control group consisting of the other New England states because of their geographic proximity to Massachusetts. An additional specification excludes states that passed more limited health care reforms during the sample period (California, Hawaii, Maine, Oregon and Vermont).

The sixth robustness check constructs a “synthetic control group” for Massachusetts, as described by Abadie et al. (2010). We first aggregate to the state-by-year level and allow the data to select the combination of the other 50 states that best matches Massachusetts on health status and the control variables during the pre-treatment years 2001-2005.²⁸ The resulting control group is 70.9% Connecticut, 11.3% Rhode Island, 8% Washington, D.C., 5.9% Utah, 3.7% California, and 0.1% Arizona. Following Fitzpatrick’s (2008) application of this method to individual data, we then multiply the weights for the individual-level observations by these shares, leaving Massachusetts fully weighted and dropping the 44 states that received a zero weight.²⁹

The next regression uses the rest of the country as the control group but excludes the year 2005. Recall from Figure 2.1 that in 2005 health in Massachusetts was below the

²⁷ When matching on pre-treatment levels, the control states are Colorado, Connecticut, District of Columbia, Maryland, Minnesota, Nebraska, New Hampshire, Utah, Vermont and Virginia. When matching on pre-treatment trends, the control states are Arkansas, California, Hawaii, Illinois, Indiana, Maine, Mississippi, Missouri, New Jersey and New York. When matching on pre-treatment coverage, the control states are Connecticut, Delaware, District of Columbia, Hawaii, Iowa, Maryland, Michigan, Pennsylvania, Rhode Island and Wisconsin. Unreported regressions used control groups of five or twenty states instead of ten; the results were similar.

²⁸ We do this using the Stata module “synth” (Abadie et al., 2011).

²⁹ In the “matching on pre-treatment levels,” “matching on pre-treatment trends,” New England, and synthetic control regressions, the number of states is 11 or fewer. Angrist and Pischke (2008) note that standard errors clustered by state are unreliable when the number of states is small. As they recommend, we instead cluster standard errors at the state-by-year level in these four regressions.

trend line, raising the question of whether the improvement in health from 2005 to 2006 could be due to a temporary negative shock in 2005 rather than the reform in 2006. The long pre-treatment period mitigates this concern by tempering the influence of 2005, but dropping 2005 addresses it more directly.³⁰

Finally, we return to the full sample but control for the potential time-varying state-level confounders unemployment rate, cigarette tax rate, physician density, and hospital density, along with linear state-specific time trends to allow for differential trends in health along unobservable dimensions.³¹ Controlling for unemployment rate and cigarette tax could be especially important given the differential impacts of the recession across states and the large cigarette tax increase passed in Massachusetts in 2009.

We present the results of these robustness checks in Columns 2 through 9 of Table 2.4. The coefficient of the interaction term $MA * During$ remains positive in all specifications, with magnitudes ranging from 0.010 to 0.022, though it loses statistical significance in some of the regressions with smaller control groups. In contrast, the interaction term $MA * After$ remains highly significant in all specifications. The magnitude of its effect is stable, as it ranges from 0.032 to 0.049 and is always within the 95% confidence interval from the baseline regression. As a result the treatment effects are also similar across specifications.

³⁰ Other unreported robustness checks experimented with the use of shorter pre-treatment periods beginning in 2002, 2003, or 2004. The results remained very similar.

³¹ We relegate the state-level control variables to a robustness check rather than using them in the main analysis because of concerns that some of them – in particular unemployment rate, physician density, and hospital density – could be endogenous to health care reform. Moreover, the four state-level controls are all individually and jointly insignificant, so the state fixed effects appear to sufficiently capture their influence on health.

Testing for Differential Pre-Treatment Trends and Delayed Effects

This section simultaneously addresses two possible concerns with the estimates from Table 2.4. First, the difference-in-differences approach assumes common counterfactual health trends between Massachusetts and the rest of the country. The robustness of the estimates to different constructions of the control group is consistent with this assumption, but conceivably health trends in Massachusetts could be so unique that no appropriate comparison group of states exists. Second, the preceding regressions do not differentiate between the short- and long-run health effects of the reform following full implementation. Since health is a capital stock accumulated through repeated investments, the improvements in health resulting from the reform could increase over time. Alternatively, the long-term uninsured might experience a pent-up demand for medical services after obtaining coverage, in which case the entire improvement in health could be reached quickly or even be temporary.

We address these issues by re-estimating equation (2.1) with a broader set of interaction terms. First, we divide the ten-year sample into five two-year periods and include interactions of the Massachusetts dummy with indicators for 2003-2004, 2005-2006, 2007-2008, and 2009-2010 (leaving 2001-2002 as the reference period). A second regression interacts Massachusetts with a full set of year dummies. These models test the common trends assumption by testing for differential trends between Massachusetts and other states in the pre-treatment period 2001-2005. If the treatment and control groups were trending similarly before the reform, then they likely would have continued to trend similarly from 2006-2010 if the reform had not occurred. The models also distinguish

between short- and long-run effects by including multiple interactions from the post-reform period.

Table 2.5 displays the coefficient estimates for the interaction terms. The regression with two-year splits shows that health trends in Massachusetts and other states were similar through the pre-treatment period, with a sizeable gap emerging in the early period following the reform's full implementation (2007-2008) that grew only slightly in the later period (2009-2010). These results are consistent with the reform having a positive causal effect on health, and with the short- and long-run effects being similar. The results from the one-year splits are broadly similar, with the exception that Massachusetts experienced a temporary negative health shock in 2005 that disappeared by 2006. At no point in the pre-treatment period was there a Massachusetts-specific health shock that lasted longer than one year, making it unlikely that the sustained improvement in health in Massachusetts from 2007-2010 would have occurred in the absence of the reform. Moreover, the regression excluding 2005 from Table 2.4 provides further evidence that the negative shock in Massachusetts in 2005 is not meaningfully influencing our conclusions.³²

³² As an alternative approach to testing the common trend assumption, in Appendix C we conduct three falsification tests restricting the sample to the pre-treatment years 2001-2005. The first considers 2001-2003 to be the “before” period and 2004-2005 the “after” period, while the second treats 2001-2002 as the “before” period and 2003-2005 as the “after” period. The third classifies 2001-2002 as the “before” period, 2003 as the “during” period, and 2004-2005 as the “after” period. None of these tests produce any evidence of differential pre-treatment trends between Massachusetts and the other states.

Testing for Endogenous Moving Patterns

The Massachusetts reform's coverage expansions likely appeal to individuals with pre-existing conditions or a higher probability of facing future illness. This section therefore addresses another possible concern: that Massachusetts attracted sicker residents after the reform, either by making them less likely to leave the state or more likely to move there. If this is the case, our estimates may understate the reform's true effect on health, as the positive causal effect would be tempered by negative selection.

We test for endogenous moving patterns by examining whether the demographic and financial profile of Massachusetts residents changed following the reform in a way that would suggest a change in the underlying propensity towards health of the state's population. We first conduct a linear regression of health status index on the individual-level control variables among the pre-treatment portion of the sample, using the coefficient estimates to predict health for the entire sample. We then estimate the influence of $MA_s * During_t$ and $MA_s * After_t$, along with the state and time fixed effects, on predicted health status. Table 2.6 reports the results. The coefficient estimates for the interaction terms are both negative, consistent with Massachusetts health care reform attracting sick individuals, but the effects are small and insignificant at the 5% level. It therefore seems unlikely that endogenous moving patterns are meaningfully attenuating the estimated impact of the reform on health.

Tests Related to Inference

This section conducts tests to help rule out the possibility that the statistical significance observed in the baseline regression is merely an artifact of underestimated standard errors. First, following Bertrand et al.'s (2004) suggestion, we compress all the available data into a state-level panel with three time periods – “before”, “during”, and “after” – and regress state average health index on $MA_s * During_t$, $MA_s * After_t$, and state and time period fixed effects. Next, we compress the data into only two cross-sectional units – Massachusetts and other states – and ten years, defining 2006 and 2007 as the “during” period and 2008 to 2010 as the “after” period. We then regress average health index on $MA_s * During_t$, $MA_s * After_t$, a Massachusetts dummy, and year fixed effects. As shown in Table 2.7, $MA_s * After_t$ remains statistically significant in both regressions despite the small sample, and the effect sizes in standard deviations (of the individual-level health index) are similar to those from Table 2.4.

In the spirit of Abadie et al. (2010), we also consider a different approach to inference and ask how likely it would be to estimate similarly large health improvements simply by picking any state at random. We re-estimate the baseline ordered probit regression with each of the other 50 states as the “treated” unit. Only two states – Oregon and Florida – had larger positive “treatment effects” than Massachusetts. The probability of obtaining as large a health improvement as that estimated for Massachusetts by chance is therefore 4%, below the standard 5% significance level.³³ Moreover, the result for

³³ We do not report the full set of results for all 50 states due to space considerations; they are available upon request.

Oregon could potentially be explained by the 2008 Medicaid expansion shown to improve self-assessed health by Finkelstein et al. (2011).

Other Health Outcomes

This section moves beyond the overall health index and explores the effect of the reform on a variety of additional health outcomes: number of days out of the past 30 not in good physical health, not in good mental health, and with health-related functional limitations; activity-limiting joint pain; BMI; minutes per week of moderate physical activity and vigorous physical activity; and smoking status. These variables were chosen because they satisfy two conditions: 1) they are strongly and significantly correlated with the overall health index in the expected direction (as shown in Appendix D), and 2) they do not rely on a doctor's diagnosis, since a diagnosis requires medical access which is endogenous to the reform.³⁴

Analyzing health outcomes beyond the overall self-assessed health index serves three purposes. First, verifying that we also observe improvements in health using a wide range of more specific (and therefore less subjective) questions increases our confidence that the reform did in fact improve objective – and not merely subjective – health. Second, examining additional outcomes sheds light on the mechanisms through which this effect occurred. For instance, obtaining health insurance can improve physical (or mental) health through increased utilization of medical services, mental health through lower

³⁴ The first condition excludes, for instance, alcoholic drinks per month, which is only weakly correlated with health and in the opposite of the expected direction. The second condition excludes BRFSS questions that ask whether a respondent has ever been diagnosed with a particular chronic condition, such as diabetes and asthma.

stress from reduced financial risk, or health behaviors through expanded access to advice and information. Third, including the health behavior-related variables BMI and smoking tests a separate prediction of economic theory: reduced financial vulnerability to health shocks from insurance coverage could cause people to take more health risks, a phenomenon known as “ex ante moral hazard” (e.g. Dave and Kaestner, 2009; Bhattacharya et al., 2011).

Days not in good physical and mental health, days with health-related limitations, and minutes of moderate and vigorous exercise per week are non-negative count variables with variances higher than the means. We therefore estimate negative binomial models for these outcomes. The conditional expectation is given by

$$E[num_{ist} | \mu_{ist}, \alpha] = \mu_{ist} \quad (2.11)$$

where num is the number of days or minutes, α is the over-dispersion coefficient, and μ is defined by

$$\mu_{ist} = \exp(\gamma_0 + \gamma_1(MA_s * During_t) + \gamma_2(MA_s * After_t) + \mathbf{X}'_{ist}\boldsymbol{\gamma}_3 + \theta_s + \rho_t) \quad (2.12)$$

The treatment effect on the treated is defined as

$$\begin{aligned} \tau_{i,MA,t} &= \exp(\gamma_0 + \gamma_2 + \mathbf{X}'_{i,MA,t}\boldsymbol{\gamma}_3 + \theta_{MA} + \rho_t) \\ &\quad - \exp(\gamma_0 + \mathbf{X}'_{i,MA,t}\boldsymbol{\gamma}_3 + \theta_{MA} + \rho_t) \end{aligned} \quad (2.13)$$

while the average treatment effect on the treated is the mean of τ among Massachusetts residents in the “after” period.

For the binary outcome variables (activity-limiting joint pain and smoking status), we estimate probit models of the form

$$\begin{aligned} \Pr(y_{ist} = 1) \\ = \Phi(\delta_0 + \delta_1(MA_s * During_t) + \delta_2(MA_s * After_t) + \mathbf{X}'_{ist}\boldsymbol{\delta}_3 + \omega_s + v_t) \end{aligned} \quad (2.14)$$

with the treatment effect on the treated being

$$\tau_{i,MA,t} = \Phi(\delta_0 + \delta_2 + \mathbf{X}'_{ist}\boldsymbol{\delta}_3 + \sigma_{MA} + v_t) - \Phi(\delta_0 + \mathbf{X}'_{ist}\boldsymbol{\delta}_3 + \omega_{MA} + v_t).^{35} \quad (2.15)$$

Body mass index is continuous, so we estimate a linear regression in which the treatment effect is simply the coefficient estimate for $MA_s * After_t$.

Some of the health-related questions were not asked in Massachusetts in certain years, necessitating restrictions to the sample. Activity-limiting joint pain and the two measures of exercise are only available in odd-numbered survey years, meaning that the “during” period spans only six months (January 2007 to June 2007). We therefore combine those six months with the rest of 2007 and 2009 and classify the two years as the “after” period, dropping the $MA_s * During_t$ interaction from those regressions. Additionally, the physical health, mental health, and health limitations variables are not available in 2002.

³⁵ We include cigarette tax as an additional covariate in the smoking regression.

Table 2.8 presents the results using the full control group of 50 states.³⁶ Health care reform in Massachusetts is associated with reductions in the number of days not in good physical health, not in good mental health, and with health-related functional limitations, as well as a lower probability of having activity-limiting joint pain. The magnitudes of these reductions range from 0.018 to 0.033 standard deviations, roughly similar to the size of the effect for the overall health status index. It therefore seems unlikely that the observed effect on the health index is driven purely by the subjectivity of the question. Moreover, these results suggest that the reform improved health more broadly than merely by reducing stress from lower financial risk.

Turning to the health-behavior related variables, the reform is associated with a 0.025 standard deviation reduction in BMI and a 0.036 standard deviation increase in moderate exercise, but no statistically detectable effect on vigorous exercise or smoking. These results suggest that expanded access to primary care improves at least some health behaviors, perhaps through information or accountability. The increase in moderate but not vigorous exercise is consistent with physician advice encouraging sedentary individuals to begin a light exercise routine, rather than encouraging those who are already active to increase or intensify their activity. The non-effect on smoking is consistent with evidence that smoking habits respond only gradually to external factors (e.g. Courtemanche, 2009), but could also reflect the health consequences of smoking already being widely-known even without physician access. Importantly, none of the

³⁶ To conserve space, we do not present the results from the full range of specifications from Table 2.4 for these other health outcomes. Unreported regressions verify that the conclusions reached are not driven by the choice of control group.

regressions provide any evidence of ex-ante moral hazard causing individuals to take more health risks after obtaining insurance.

The final column of Table 2.8 presents the results using as the dependent variable a “cardinalized overall health status index” equal to the predicted outcome from a regression of the health index on the six most plausibly objective health outcomes: functional limitations, joint pain, BMI, moderate exercise, vigorous exercise, and smoking ($R^2 = 0.27$).³⁷ This approach is advocated by Ziebarth et al. (2010) and others as a way to handle reporting heterogeneity in self-assessed health. The impact of $MA_s * After_t$ remain positive and significant, and the effect size in standard deviations is similar to those from Table 2.4. This provides further evidence that our conclusions are not merely driven by subjectivity.

Heterogeneity

We next return to the actual overall health status index and examine heterogeneity in the effect of Massachusetts health care reform on the bases of gender, age, race, and income. Kolstad and Kowalski (2010) found the largest coverage expansions among men, young adults, minorities, and those with low incomes. However, different effects on coverage do not necessarily translate to different effects on health, as the impacts of coverage on health could also be heterogeneous. We consider the following subsamples: women; men; ages 18-34, 35-44, 45-54, 55-64, 65-74, and 75 and older; whites; blacks; Hispanics; other race; and household incomes below \$25,000,

³⁷ We also considered dropping health limitations from the set of variables used to make the prediction, or using all eight alternate health outcomes to make the prediction. The results were virtually identical.

between \$25,000 and \$75,000, and above \$75,000. We choose these income splits in order to loosely align with the provisions of the reform, which specify that health insurance be free up to 150% FPL (\$23,050 for a family of four) and subsidized up to 300% (\$69,150 for a family of four).³⁸ We estimate the baseline ordered probit model for all subsamples, with one exception. The baseline model gives an implausibly large magnitude for the 75 and older subsample, which upon further investigation appears to be driven by differential pre-treatment trends between Massachusetts and non-Massachusetts residents of that age group. We therefore include linear state-specific trends for that subsample.³⁹

Table 2.9 reports the results for the gender and age subsamples. The impact on health is positive and significant for both women and men but stronger for women. Stratifying by age, the effect is largest among the near-elderly aged 55-64, second largest among those 45-54, smaller among the two groups below 45, and smaller still among the two elderly groups. Our finding that the effect of the reform diminishes dramatically at age 65 is not surprising since individuals eligible for Medicare cannot purchase insurance through the Connector (Blue Cross Blue Shield of Massachusetts, 2006). It is interesting, though, that we still observe some evidence of health improvements among the elderly despite Medicare. Only those seniors who have paid Medicare taxes for at least ten years (or whose spouse has done so) are eligible for free Medicare Part A (Johnson-Lans, 2005),

³⁸ 2012 federal poverty lines are available at coverageforall.org/pdf/FHCE_FedPovertyLevel.pdf, accessed 6/26/12. Since the BRFSS only reports income categories and lacks comprehensive information about household size, lining up the categories to exactly match 150% and 300% of the poverty line is not possible.

³⁹ Recall that differential pre-treatment trends did not appear to be an issue for the full sample, and including state-specific trends for the full sample did not meaningfully affect the results. This suggests that the baseline estimator without state trends is still appropriate for the full sample, even if it is not for the 75 and over age group.

and presumably those seniors ineligible for Medicare could purchase community-rated insurance through the Connector. Indeed, in our data the reform increases the coverage rate of the elderly by a statistically significant 0.3 percentage points, an effect similar to that found by Kolstad and Kowalski (2010) using the National Inpatient Sample. Moreover, seniors could be affected by system-wide changes in the delivery of health care following the reform, such as reduced crowds in emergency rooms or the improvements in some dimensions of quality of care noted by Kolstad and Kowalski (2010).

Table 2.10 stratifies by race and income. Chapter 58 improved health across all racial subgroups, but the effect was largest for blacks and those of a race besides white, black, or Hispanic. A back-of-the-envelope calculation provides a ballpark estimate of the extent to which the reform reduced the health disparity between blacks and whites. In Massachusetts in the “before” period, the mean health status indices of blacks and whites were 2.553 and 2.786, respectively, for a difference of 0.233. The treatment effects imply changes in the health status indices of blacks and whites of 0.091 and 0.036, for a difference of 0.055. We therefore estimate that the reform reduced black-white health disparities by 23.6%. Stratifying by income, the reform improved the health of all three income groups but had the largest effect amongst those with incomes below \$25,000 for whom insurance premiums are heavily or fully subsidized.

Instrumental Variables

We close the empirical analysis by using $MA_s * During_t$ and $MA_s * After_t$ as instruments to estimate the impact of having insurance coverage on health. This

instrumental variables approach requires stricter assumptions than the reduced-form model, as the reform must only impact health along the extensive margin of insurance coverage, conditional on the controls. This assumption would be violated if the reform also influenced health through the intensive margin of coverage, for instance by causing some individuals to switch from high-deductible catastrophic coverage to more comprehensive coverage available through the Connector. This assumption would also be violated if the reform affected the health of those who did not switch insurance plans through system-wide changes to health care delivery or peer effects. Despite these caveats, the instrumental variables analysis is useful because it estimates the magnitude of the impact of insurance on health that would be necessary for the extensive margin to be the only channel through which the reform influenced health. If the magnitude is implausibly large, then other mechanisms must play a role as well. Since the assumption that the entire effect on health occurs through the extensive margin of coverage is unlikely to hold for Medicare beneficiaries, we exclude seniors from the analysis in this section.

The first stage predicts insurance coverage using the following linear probability model:

$$ins_{ist} = \alpha_0 + \alpha_1(MA_s * During_t) + \alpha_2(MA_s * After_t) + \mathbf{X}'_{ist}\boldsymbol{\alpha}_3 + \zeta_s + \eta_t + u_{ist} \quad (2.16)$$

where *ins* is a dummy variable equal to 1 if the person reported having any health insurance coverage. Because of the non-linearity of the second stage, we utilize a two-stage residual inclusion (2SRI) approach in which the residual from the first-stage

regression is included as an additional regressor in the second stage. Terza et al. (2008) show that in non-linear contexts 2SRI gives consistent coefficient estimates, while traditional two stage least squares does not. The second stage is modeled as an ordered probit and the probabilities of being in each of the five health states are given by,

$$\Pr(y_{ist} = 0) = \Phi(\lambda_1 - \pi_1 ins_{ist} - \mathbf{X}'_{ist}\boldsymbol{\pi}_2 - \pi_3 \hat{u}_{ist} - \sigma_s - \varphi_t) \quad (2.17)$$

$$\Pr(y_{ist} = k) = \Phi(\lambda_j - \pi_1 ins_{ist} - \mathbf{X}'_{ist}\boldsymbol{\pi}_2 - \pi_3 \hat{u}_{ist} - \sigma_s - \varphi_t) - \Phi(\lambda_{j-1} - \pi_1 ins_{ist} - \mathbf{X}'_{ist}\boldsymbol{\pi}_2 - \pi_3 \hat{u}_{ist} - \sigma_s - \varphi_t), \forall j \in (2,3,4) \quad (2.18)$$

$$\Pr(y_{ist} = 4) = 1 - \Phi(\lambda_4 - \pi_1 ins_{ist} - \mathbf{X}'_{ist}\boldsymbol{\pi}_2 - \pi_3 \hat{u}_{ist} - \sigma_s - \varphi_t) \quad (2.19)$$

where \hat{u}_{ist} is the first-stage residual. The effect of health insurance on the probability of being in health state j is

$$\Delta p_j = \Pr(y_{ist} = j | ins_{ist} = 1) - \Pr(y_{ist} = j | ins_{ist} = 0) \quad (2.20)$$

The asymptotic standard errors of these probabilities and the standard errors for the second stage estimates were calculated following Terza (2011).⁴⁰ Equation (2.19) represents the “local average treatment effect” of insurance among those who obtained coverage as a result of the reform, and is subject to the usual caveat regarding generalizability.

Table 2.11 reports the coefficient estimates of interest from the first and second stage regressions for the full sample, along with the estimated impacts of insurance on the

⁴⁰ Mata code is available upon request.

health state probabilities. The first stage estimates an increase in the coverage rate of 1.9 percentage points in the “during” period and 5.4 percentage points in the “after” period. The F statistic from a test of the joint significance of $MA_s * During_t$ and $MA_s * After_t$ is large, suggesting the instruments are sufficiently strong. Turning to the second stage, obtaining insurance leads to a positive and statistically significant improvement in health. The first-stage residual is significant and negatively associated with health, providing evidence that an OLS estimator would suffer from a downward bias. Insurance is estimated to reduce the probabilities of being in poor, fair, and good health by 6.2, 9.8, and 8.5 percentage points, while increasing the probabilities of being in very good and excellent health by 8.5 and 16 percentage points. The overall effect of insurance on the health status index, encompassing changes in all five probabilities, is 0.585 of the sample standard deviation.

These effects are strikingly large, but assessing their plausibility requires a comparison to other estimates from the literature. Finkelstein et al. (2011) employ the cleanest research design to date among studies of the impact of insurance on self-assessed health: a randomized intervention in Oregon granting Medicaid eligibility to a subset of the uninsured. They estimate that Medicaid enrollment increases the probability of being in good, very good, or excellent health by 13.3 percentage points. The sum of our estimated effects on the probabilities of being in those three health states is a similar 16 percentage points. The results from the two papers are not directly comparable given the differences in populations, but this similarity suggests that it is conceivable that the reform’s entire effect on the self-assessed health of the non-elderly could have occurred

through the extensive margin of coverage. Future research should more directly investigate the roles of other potential channels.

We also conduct instrumental variables analyses for the gender, race, age, and income subgroups, allowing us to assess whether the heterogeneity in the reform's effect on health observed in the Heterogeneity Section comes from heterogeneity in the effect *on* coverage or the effect *of* coverage. Appendices E and F report the results. The coverage expansions are larger for men than women, but women have greater health gains from coverage, explaining the greater net effect of the reform for women. Among the age subsamples, those under 35 years old have the largest gains in coverage, but also the smallest health improvements from obtaining coverage. Of the non-elderly age groups, 55-64 year olds have the smallest effect of the reform on coverage but the largest effect of coverage on health. Stratifying by race shows that coverage rates increase the most for non-black minorities but that the health effects of coverage are the largest for blacks. Finally, the coverage expansions are the largest for the low-income group, second largest for the middle-income group, and smallest for those with high incomes. However, the effect of coverage on health is the strongest for the high income group.

Conclusion

This paper examined the effect of health care reform in Massachusetts on self-assessed health using data from the Behavioral Risk Factor Surveillance System (BRFSS). An ordered probit difference-in-differences analysis showed that the reform increased the probability of individuals reporting excellent or very good health while reducing their probability of reporting good, fair, or poor health. These results were robust to alternative

constructions of the control group and the addition of state-level covariates. We did not find evidence that the estimates were meaningfully impacted by differential pre-treatment trends or endogenous moving patterns. Next, we examined a number of more specific health outcomes and found improvements in physical health, mental health, functional limitations, joint disorders, body mass index, and moderate physical activity. Testing for heterogeneity revealed that women, minorities, near-elderly adults, and those with incomes low enough to qualify for the law's subsidies experienced the largest gains in health as a result of the reform. Finally, we embedded the reform in an instrumental variables framework and estimated a large positive impact of obtaining health insurance on health.

Perhaps the clearest limitation of our analysis is that all our health outcomes were self-reported. Our finding of similar results across a range of health outcomes with varying degrees of subjectivity increases our confidence that our findings largely represent “real” changes in physical/mental health. However, we cannot rule out the possibility that some of the observed improvement in health could merely be due to a “warm glow” from acquiring health insurance. To underscore this point, recall that our estimated effects of insurance on self-assessed health are a similar magnitude to those of Finkelstein et al. (2011), and they found that a sizeable portion of the reported health improvements following the Oregon experiment occurred prior to measurable changes in overall health care utilization. Obtaining insurance coverage can reduce stress, which can directly improve numerous aspects of health even without any additional medical care

being utilized, but Finkelstein et al. (2011) do raise the question of what the estimated improvements in self-assessed health are capturing.

We argue that Finkelstein et al.'s (2011) finding regarding timing does not automatically apply to our context for several reasons. First, their data only tracked individuals for a year after the intervention, while we have 4½ years of data after first of the newly-insured in Massachusetts obtained coverage and 3½ years after all major facets of the reform took effect. If a “warm glow” from acquiring insurance was driving the effect, we would have expected the reported health benefits from the reform to diminish over time, but as Table 2.5 shows this was not the case. Second, other studies have documented changes in health care utilization in Massachusetts at around the same time as we observed health improvements (Kowalski and Kolstand, 2010; Miller, 2011a). Next, the newly insured in the Oregon experiment were winners of a random lottery, which could lead to a stronger “warm glow” than simply acquiring health insurance from a statewide intervention like the reform in Massachusetts. Accordingly, we consistently find that the effects on health were small at best in the “during” period, which includes nine months after those with incomes below 100% FPL became eligible for free coverage. We therefore do not seem to observe the immediate spike in self-assessed health seen in the Oregon experiment. Nonetheless, as the necessary data become available it will be important to evaluate the impact of the Massachusetts reform on unambiguously objective measures of health such as mortality.

Another natural question is the degree to which our results from Massachusetts can serve as projections for the Affordable Care Act. The general strategies for obtaining

nearly universal coverage in both the Massachusetts and federal laws involved the same three-pronged approach of non-group insurance market reforms, subsidies, and mandates, suggesting that the health effects should be broadly similar. However, the federal legislation included additional cost-cutting measures such as Medicare cuts that could potentially mitigate the gains in health from the coverage expansions. On the other hand, baseline uninsured rates were unusually low in Massachusetts, so the coverage expansions – and corresponding health improvements – from the Affordable Care Act could potentially be greater. Of course, larger coverage expansions may mean higher costs, and costs should be weighed against benefits when evaluating the welfare implications of reform.

Tables and Figures

Figure 2.1 – Changes in Health Status Index 2001-2010

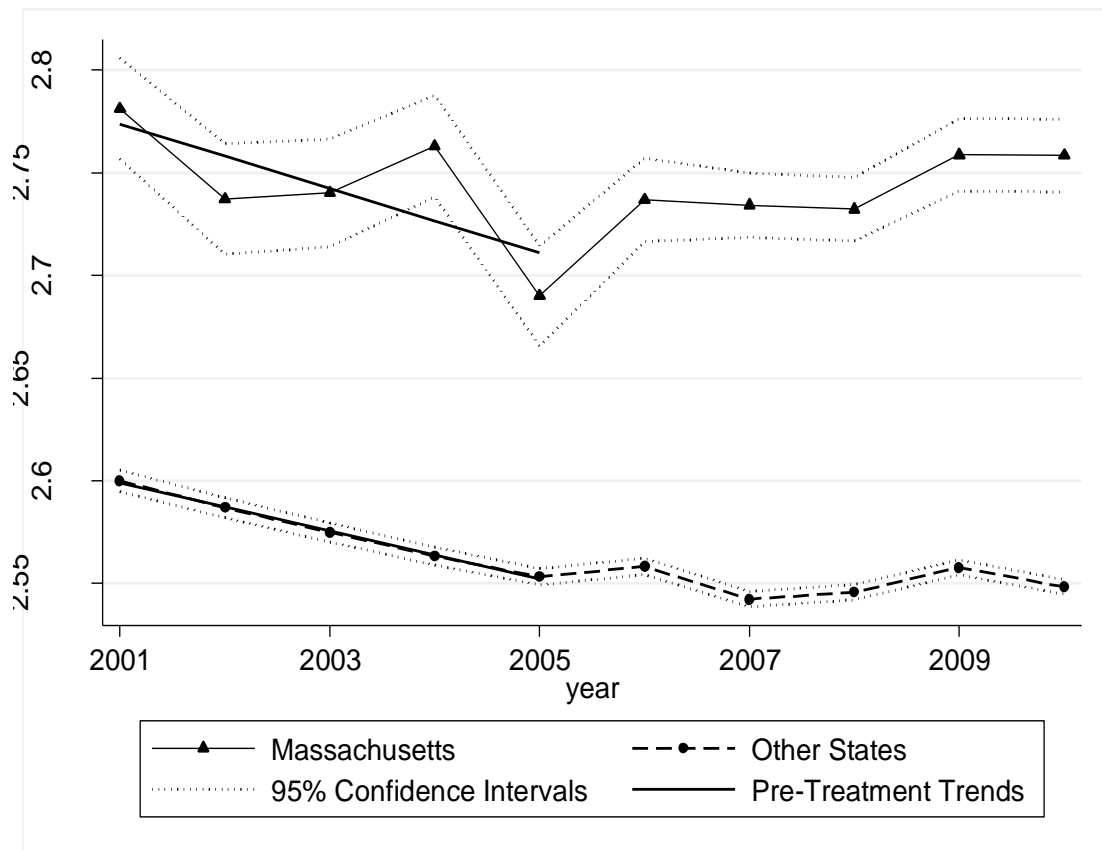


Table 2.1 – Similarities and Differences between the Massachusetts Reform and the National Reform (ACA)

Domain	Massachusetts reform	National reform (ACA)
Modification of existing insurance markets	<ul style="list-style-type: none"> - No pre-existing condition exclusions (since 1996). - Community rated premiums that can only vary by age and smoking status (in place since 1996). - Minimum standards for policies, including essential benefits and maximum out of pocket expenditures. - Creation of a state health insurance exchange where insurance companies compete to offer three regulated levels of coverage to small employers and individuals. - Young adults must be allowed coverage on their parents' plans for up to two years after they are no longer dependents or until their 26th birthday. 	<ul style="list-style-type: none"> - No pre-existing condition exclusions. - Community rated premiums that can only vary by age and smoking status. - Minimum standards for policies, including essential benefits and maximum out of pocket expenditures. - States must create a health insurance exchange where insurance companies compete to offer four regulated levels of coverage to small employers and individuals. States are able to join multistate exchanges. - Young adults must be allowed coverage on their parents' plans until their 26th birthday.
Mandates	<ul style="list-style-type: none"> - Individuals are required to purchase coverage if affordable, (based on income and family size) or pay a penalty of no more than 50% of the insurance premium of the lowest-cost insurance exchange plan for which they are eligible. - Employers with more than 10 full time employees (FTE) are required to offer policies with minimum standard or pay a penalty of up to \$295 annually per FTE. 	<ul style="list-style-type: none"> - Individuals are required to purchase coverage if it costs no more than 8% of income, or pay a penalty of the greater of 2.5 percent of taxable income or \$695. - Employers with 50 employees or more are required to offer policies with minimum standard or pay penalties that range from \$2,000-\$3,000 per FTE.
Medicaid expansions and subsidies	<ul style="list-style-type: none"> - Medicaid expansions for children with household incomes up to 300% of the poverty line (FPL), for long-term unemployed up to 100% FPL, and for people with HIV up to 200% FPL. - Free coverage for all adults below 150% FPL. Sliding scale of subsidies for adults up to 300% FPL. 	<ul style="list-style-type: none"> - Medicaid expansions to all individuals with incomes below 133% FPL. - Sliding scale of tax credits for people up to 400% FPL. - Tax credits for employers with 25 or fewer employees and average annual wages less than \$50,000 for offering coverage.
Financing	<ul style="list-style-type: none"> - Redirection of federal funding to safety net providers. - Redirection of the state uncompensated care pool, a mechanism through which hospitals were able to bill 	<ul style="list-style-type: none"> - Reduction of Medicare reimbursements. - Increase in the Medicare payroll tax and extension of this tax to capital income for singles (families) with

the state the cost of treating low-income patients.	incomes more than \$200,000 (\$250,000).
- Individual and employer penalties.	- Individual and employer penalties.
- One-time assessment to health care providers and insurers.	- Taxes on insurers, pharmaceutical companies, and medical device manufactures.
- Since 2009, a \$1 per pack cigarette tax.	- Excise taxes on high-cost insurance plans (“Cadillac tax”).

Sources: Gruber (2011a, 2008b) and Harrington (2010).

Table 2.2 – Pre-Treatment Means of Health Variables

Variable	MA (n=35,990)	Other States (n=1,177,056)	Difference
Any health insurance coverage	0.911	0.848	-0.063***
Overall health; 0 (poor) to 4 (excellent)	2.743	2.575	0.168***
Poor health	0.030	0.041	-0.011***
Fair health	0.089	0.112	-0.023***
Good health	0.261	0.294	-0.034***
Very good health	0.351	0.336	0.015***
Excellent health	0.270	0.212	0.054***
Days not in good physical health (of last 30) ⁺⁺	3.271	3.479	-0.207***
Days not in good mental health (of last 30) ⁺⁺	3.307	3.437	-0.130*
Days with health limitations (of last 30) ⁺⁺	1.916	2.080	-0.164***
Activity-limiting joint problems ⁺	0.123	0.133	-0.009**
Body mass index	26.319	26.992	-0.673***
Minutes of moderate exercise per day ⁺	57.658	58.788	-1.130
Minutes of vigorous exercise per day ⁺	40.065	38.830	1.235
Currently smokes cigarettes	0.192	0.224	-0.032***

Notes: *** indicates difference between Massachusetts and other states is significant at the 0.1% level; ** 1% level; * 5% level. Observations are weighted using the BRFSS sampling weights. + indicates variable from only odd-numbered survey years. ++ indicates variable from all years except 2002. Standard errors are available on request.

Table 2.3 – Pre-Treatment Means of Control Variables

Variable	Massachusetts (n=35,990)	Other States (n=1,177,056)	Difference
Age 18 to 24	0.114	0.121	-0.007*
Age 25 to 29	0.083	0.089	-0.006**
Age 30 to 34	0.107	0.105	0.002
Age 35 to 39	0.105	0.104	0.001
Age 40 to 44	0.116	0.112	0.004
Age 45 to 49	0.100	0.101	-0.001
Age 50 to 54	0.089	0.091	-0.003
Age 55 to 59	0.073	0.073	0.000
Age 60 to 64	0.056	0.056	0.000
Age 65 to 69	0.045	0.047	-0.002
Age 70 to 74	0.038	0.040	0.002
Age 75 to 79	0.039	0.034	0.005***
Age 80 or older	0.035	0.030	0.005***
Female	0.510	0.502	0.008*
Married	0.571	0.598	-0.028***
Race is non-Hispanic white	0.846	0.709	0.137***
Race is non-Hispanic black	0.034	0.098	-0.063***
Race is Hispanic	0.113	0.178	-0.065***
Race is neither black nor white nor Hispanic	0.006	0.015	-0.009***
Income less than \$10,000	0.037	0.055	-0.018***
Income \$10,000 to \$15,000	0.039	0.056	-0.017***
Income \$15,000 to \$20,000	0.057	0.079	-0.022***
Income \$20,000 to \$25,000	0.077	0.096	-0.019***
Income \$25,000 to \$35,000	0.108	0.139	-0.031***
Income \$35,000 to \$50,000	0.149	0.172	-0.023***
Income \$50,000 to \$75,000	0.188	0.174	0.014***
Income \$75,000 or more	0.345	0.228	0.117***
Less than a high school degree	0.071	0.114	-0.043***
High school degree but no college	0.251	0.299	-0.048***
Some college but not four-year degree	0.242	0.273	-0.031***
College graduate	0.436	0.314	0.121***
Currently pregnant	0.011	0.012	-0.001
State unemployment rate	4.979	5.435	-0.456***
State cigarette tax (20010 \$)	1.485	0.820	0.665***
State physician density (per 10,000 residents)	436.247	256.945	179.302***
State hospital density (per 10,000 residents)	1.208	1.701	-0.493***

Notes: *** indicates difference between Massachusetts and other states is significant at the 0.1% level; ** 1% level; * 5% level. Observations are weighted using the BRFSS sampling weights. Standard errors are available on request.

Table 2.4 – Difference-in-Differences Ordered Probit Regressions

	Dependent Variable: Overall Health								
	Full Sample	Match on Pre-Tx. Level	Match on Pre-Tx. Trend	Match on Pre-Tx. Coverage	New England	Drop CA, HI, ME, OR, VT	Synthetic Control Group	Drop 2005	Add State Controls/Trends
Coefficient Estimates of Interest									
MA*Du-ring	0.017 (0.006)**	0.013 (0.016)	0.022 (0.015)	0.013 (0.015)	0.015 (0.009)	0.016 (0.007)***	0.013 (0.007)*	0.010 (0.014)	0.022 (0.010)*
MA*After	0.039 (0.006)***	0.049 (0.010)***	0.037 (0.008)***	0.046 (0.010)***	0.049 (0.007)***	0.038 (0.006)***	0.044 (0.008)***	0.032 (0.007)***	0.049 (0.010)***
Average Treatment Effects on Treated (After Period)									
P(Poor)	-0.002 (0.0003)***	-0.003 (0.0006)***	-0.002 (0.0005)***	-0.002 (0.0006)***	-0.003 (0.0004)***	-0.002 (0.0004)***	-0.002 (0.0005)***	-0.002 (0.0003)***	-0.003 (0.0006)***
P(Fair)	-0.005 (0.0007)***	-0.006 (0.001)***	-0.004 (0.0009)***	-0.005 (0.001)***	-0.006 (0.0009)***	-0.004 (0.0007)***	-0.005 (0.001)***	-0.004 (0.0007)***	-0.006 (0.001)***
P(Good)	-0.007 (0.0009)***	-0.008 (0.002)***	-0.006 (0.001)***	-0.008 (0.002)***	-0.008 (0.001)***	-0.007 (0.001)***	-0.007 (0.001)***	-0.005 (0.001)***	-0.008 (0.002)***
P(Very Good)	0.002 (0.0003)***	0.002 (0.0006)***	0.002 (0.0004)***	0.002 (0.0006)***	0.002 (0.0004)***	0.002 (0.0003)***	0.002 (0.0004)***	0.001 (0.0003)***	0.002 (0.0006)***
P(Excellent)	0.012 (0.002)***	0.014 (0.003)***	0.011 (0.002)***	0.013 (0.003)***	0.014 (0.002)***	0.011 (0.002)***	0.013 (0.002)***	0.010 (0.002)***	0.015 (0.003)***

Overall Effect in Std.Dev.	0.033	0.037	0.032	0.035	0.037	0.032	0.037	0.027	0.041
#Control States	50	10	10	10	5	45	6	50	50
Obs.	2,879,296	633,979	643,302	578,530	340,592	2,664,194	390,453	2,582,055	2,879,296

Notes: Standard errors, heteroskedasticity-robust and clustered by state, are in parentheses. In columns 2-5 and 7, standard errors are clustered at the state*year level rather than state because of the small number of states. *** indicates statistically significant at the 0.1% level; ** 1% level; * 5 % level. All regressions include the individual-level control variables, state fixed effects, and fixed effects for each month in each year. Observations are weighted using the BRFSS sampling weights.

Table 2.5 – Testing for Differential Pre-Treatment Trends and Delayed Effects

Dependent Variable: Overall Health		
	2-Year Splits	1-Year Splits
MA*2003 to 2004	0.004 (0.007)	--
MA*2005 to 2006	-0.014 (0.007)	--
MA*2007 to 2008	0.032 (0.005)***	--
MA*2009 to 2010	0.039 (0.007)***	--
MA*2002	--	-0.013 (0.010)
MA*2003	--	-0.014 (0.009)
MA*2004	--	0.010 (0.008)
MA*2005	--	-0.039 (0.007)***
MA*2006	--	-0.0009 (0.009)
MA*2007	--	0.026 (0.008)**
MA*2008	--	0.024 (0.008)***
MA*2009	--	0.036 (0.010)***
MA*2010	--	0.028 (0.009)**
Observations	2,879,296	2,879,296

Notes: Coefficient estimates are shown; average treatment effects on the treated are available upon request. Standard errors, heteroskedasticity-robust and clustered by state, are in parentheses. *** indicates statistically significant at the 0.1% level; ** 1% level; * 5% level. All regressions include the individual-level control variables, state fixed effects, and fixed effects for each month in each year. The control group consists of all 50 other states. Observations are weighted using the BRFSS sampling weights.

Table 2.6 – Testing for Endogenous Moving Patterns

Dependent Variable: Predicted Health Status	
	Coefficient Estimates
MA*During	-0.008 (0.004)
MA*After	-0.008 (0.008)
Effect in Standard Deviations (After Period)	-0.016
Observations	2,888,559

Notes: The coefficient estimates are equal to the treatment effects because the model is linear. Standard errors, heteroskedasticity-robust and clustered by state, are in parentheses. *** indicates statistically significant at the 0.1% level; ** 1% level; * 5 % level. The regression includes state fixed effects, and fixed effects for each month in each year. The control group consists of all 50 other states. Observations are weighted using the BRFSS sampling weights.

Table 2.7 – Regressions with Aggregated Data

Dependent Variable: Average Health Status		
	State-Level with Three Time Periods	Annual with Two Cross-Sectional Units (MA and not MA)
MA*During	0.011 (0.006)	0.018 (0.013)
MA*After	0.029 (0.005)***	0.032 (0.014)*
Effect in Standard Deviations (After Period)	0.027	0.030
Observations	153	20

Notes: The coefficient estimates are equal to the treatment effects because the model is linear. Heteroskedasticity-robust standard errors (clustered by state in the first column) are in parentheses. *** indicates statistically significant at the 0.1% level; ** 1% level; * 5 % level. The first regression includes state fixed effects and dummies for the during and after periods; the second regression includes year fixed effects and a dummy for MA. The control group consists of all 50 other states in the first regression, and one group consisting of all individuals from the 50 other states in the second regression. Observations are weighted using the BRFSS sampling weights when aggregating.

Table 2.8 – Regression Results for Other Health Outcomes

Dependent Variable:	Days not in Good Physical Health	Days not in Good Mental Health	Days with Health Limitations	Activity-Limiting Joint Pain	BMI	Minutes of Moderate Exercise	Minutes of Vigorous Exercise	Smoker	Cardinalized Overall Health
MA*After	-0.079 (0.011)***	-0.051 (0.012)***	-0.065 (0.013)***	-0.036 (0.010)***	-0.143 (0.047)**	0.039 (0.018)**	-0.002 (0.018)	0.006 (0.007)	0.013 (0.004)***
ATE on Treated	-0.255 (0.037)***	-0.165 (0.041)***	-0.128 (0.028)***	-0.006 (0.002)***	-0.143 (0.047)**+	2.026 (0.912)*	-0.079 (0.608)	0.001 (0.002)	0.013 (0.004)***+
Effect in Std. Deviations	-0.033	-0.022	-0.022	-0.018	-0.025	0.036	-0.001	0.002	0.027
Observations	2,642,885	2,649,994	2,663,473	1,333,179	2,794,388	1,217,299	1,217,299	2,878,751	1,122,083

Notes: + indicates the treatment effect and coefficient estimate are equal because the model is linear. Standard errors, heteroskedasticity-robust and clustered by state, are in parentheses. *** indicates statistically significant at the 0.1% level; ** 1% level; * 5 % level. All regressions include the individual-level control variables, state fixed effects, and fixed effects for each month in each year. MA*During is also included for all outcomes except joint pain, exercise, and cardinalized health, which are not available in odd-numbered survey years. The control group consists of all 50 other states. Observations are weighted using the BRFSS sampling weights.

Table 2.9 – Heterogeneity in the Effect on Health by Gender and Age

Dependent Variable: Overall Health								
	Gender		Age					
	Women	Men	18-34	35-44	45-54	55-64	65-74	75+
MA*After	0.046 (0.006)***	0.029 (0.006)***	0.023 (0.010)*	0.021 (0.008)**	0.036 (0.007)***	0.060 (0.011)***	0.019 (0.006)**	0.015 (0.020)
Average Treatment Effects on Treated								
P(Poor)	-0.003 (0.0003)***	-0.002 (0.0004)***	-0.0004 (0.0002)*	-0.0007 (0.0002)**	-0.002 (0.0004)***	-0.005 (0.0009)***	-0.002 (0.0006)***	-0.002 (0.003)
P(Fair)	-0.005 (0.0007)***	-0.004 (0.0008)***	-0.002 (0.0009)*	-0.002 (0.0008)**	-0.004 (0.0008)***	-0.008 (0.001)***	-0.003 (0.0009)***	-0.003 (0.003)
P(Good)	-0.008 (0.001)***	-0.005 (0.001)***	-0.005 (0.002)*	-0.004 (0.002)**	-0.006 (0.001)***	-0.009 (0.002)***	-0.002 (0.0007)***	-0.0007 (0.0009)
P(Very Good)	0.002 (0.0003)***	0.001 (0.0003)***	0.0001 (-0.0001)	0.00005 (-0.0001)	0.0009 (0.0002)***	0.004 (0.0009)***	0.002 (0.0007)***	0.003 (0.004)
P(Excellent)	0.014 (0.002)***	0.009 (0.002)***	0.008 (0.003)*	0.007 (0.003)**	0.011 (0.002)***	0.017 (0.003)***	0.005 (0.001)***	0.003 (0.004)
Overall Effect in Std. Dev.	0.039	0.024	0.022	0.018	0.029	0.049	0.017	0.015
Observations	1,733,131	1,146,165	485,376	512,575	614,489	563,405	398,264	305,187

Notes: Standard errors, heteroskedasticity-robust and clustered by state, are in parentheses. *** indicates statistically significant at the 0.1% level; ** 1% level; * 5 % level. All regressions include MA*During, the individual-level control variables, state fixed effects, and fixed effects for each month in each year. For reasons discussed in the text, the 75+ regression also includes state-specific linear trends. The control group consists of all 50 other states. Observations are weighted using the BRFSS sampling weights.

Table 2.10 – Heterogeneity in the Effect on Health by Race and Income

Dependent Variable: Overall Health							
	Race				Household Income		
	White	Black	Hispanic	Other	<\$25,000	\$25,000- \$75,000	>\$75,000
MA*After	0.036 (0.005)***	0.091 (0.012)***	0.041 (0.016)*	0.081 (0.021)***	0.061 (0.007)***	0.033 (0.007)***	0.021 (0.008)**
Average Treatment Effects on Treated							
P(Poor)	-0.002 (0.0003)***	-0.006 (0.001)***	-0.003 (0.001)*	-0.009 (0.002)***	-0.009 (0.001)***	-0.002 (0.0004)***	-0.0004 (0.0001)***
P(Fair)	-0.004 (0.0005)***	-0.013 (0.002)***	-0.006 (0.003)*	-0.011 (0.003)***	-0.011 (0.001)***	-0.004 (0.0009)***	-0.001 (0.0006)***
P(Good)	-0.006 (0.0009)***	-0.013 (0.002)***	-0.005 (0.002)**	-0.009 (0.002)***	-0.001 (0.0001)***	-0.006 (0.001)***	-0.005 (0.002)*
P(Very Good)	0.001 (0.0002)***	0.008 (0.001)***	0.003 (0.001)*	0.009 (0.003)***	0.009 (0.001)***	0.003 (0.0007)***	-0.001 (0.0004)***
P(Excellent)	0.011 (0.001)***	0.024 (0.003)***	0.011 (0.004)**	0.020 (0.005)***	0.012 (0.001)***	0.009 (0.002)***	0.008 (0.003)*
Overall Effect in Std. Dev.	0.030	0.078	0.033	0.07	0.055	0.029	0.020
Observations	2,320,271	222,581	287,895	48,549	842,088	1,346,946	690,262

Notes: Standard errors, heteroskedasticity-robust and clustered by state, are in parentheses. *** indicates statistically significant at the 0.1% level; ** 1% level; * 5 % level. All regressions include MA*During, the individual-level control variables, state fixed effects, and fixed effects for each month in each year. The control group consists of all 50 other states. Observations are weighted using the BRFSS sampling weights.

Table 2.11 – Instrumental Variables

First Stage: Any Insurance Coverage	
Coefficient Estimates	
MA*During	0.019 (0.002)***
MA*After	0.054 (0.003)***
1 st Stage F Statistic	171.42
Second Stage: Overall Health	
Coefficient Estimates	
Insurance	0.688 (0.112)***
1 st Stage Residual	-0.663 (0.112)***
Local Average Treatment Effects	
P(Poor)	-0.062 (0.011)***
P(Fair)	-0.098 (0.015)***
P(Good)	-0.085 (0.013)***
P(Very Good)	0.085 (0.013)***
P(Excellent)	0.16 (0.026)***
Overall Effect in Standard Deviations	0.585
Observations	2,172,797

Notes: A linear probability model is estimated in the first stage so the coefficient estimate equals the treatment effect. Standard errors, heteroskedasticity-robust and clustered by state, are in parentheses. ***Indicates statistically significant at the 0.1% level; ** 1% level; * 5 % level. All regressions include MA*During, the individual-level control variables, state fixed effects, and fixed effects for each month in each year. The control group consists of all 50 other states. Observations are weighted using the BRFSS sampling weights.

CHAPTER III

FOOD ASSISTANCE AND FAMILY ROUTINES IN THREE AMERICAN CITIES

(Co-authored with David Ribar)

Abstract

The Supplemental Nutrition Assistance Program, the National School Lunch Program, the School Breakfast Program, and the Special Supplemental Nutrition Program for Women, Infants, and Children all share the fundamental goal of helping needy and vulnerable people in the U.S. obtain access to nutritious foods that they might not otherwise be able to afford. However, the programs also have other objectives, such as improving recipients' health, reducing household food insecurity, and furthering children's development and school performance. To investigate these broader impacts, we examine the relationship between participation in public and private food assistance programs and a number of family processes, including family routines (including shared meals) and time use. We examine these relationships using the data from the Three-City Study, a longitudinal survey of low-income families with children to estimate multivariate longitudinal regression models that incorporate fixed effects controls that account for unobservable characteristics. Estimates from these models indicate that

SNAP participation is negatively associated with homework routines. WIC participation on the other hand, is positively associated with family routines in general and with dinner do not yield statistically significant associations between school meals and family routines.

Introduction

The major food assistance programs in the United States—the Supplemental Nutrition Assistance Program (SNAP), the National School Lunch Program (NSLP), the School Breakfast Program (SBP), and the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC)—all share the fundamental goal of helping needy families and individuals obtain access to nutritious foods that they might not otherwise be able to afford. By improving nutrition and diets, the programs also are intended to advance other goals, such as improving recipients’ health, reducing household food insecurity, and furthering children’s development and school performance. Evidence from numerous studies strongly indicates that the programs increase household food consumption and improve dietary intakes (see, e.g., Fox et al. 2004). However, the evidence regarding health, food insecurity, and schooling effects is equivocal and includes many null findings as well as some unexpected associations, such as food assistance being related to higher rates of obesity (e.g., Baum 2011, Meyerhoefer & Pylypchuk 2008), greater food insecurity (e.g., Nord et al. 2010), and even lower school attendance (Peterson et al. 2004, Ribar & Haldeman 2011). Although some of these results are likely artifacts of selection bias and other statistical issues (see, e.g., DePolt et al. 2009, Nord & Golla 2009), other results appear to be more robust.

The negative findings might be explained by unintended and largely unexamined effects of food assistance on family routines and activities. Consider one such routine—the family eating breakfast together. Merten et al. (2009) found that irregular breakfast consumption, a predictor of obesity, was higher among adolescents whose parents were not present at breakfast time. Participation in different types of food assistance programs could affect the frequency of shared breakfast times. On the one hand, participation in the SNAP could increase the availability of food and especially of food that might need additional preparation and thereby contribute to more shared household breakfasts. On the other hand, participation in the SBP could reduce household breakfasts. Also, if SNAP participation leads to more variable food availability (plentiful food early when monthly benefits are issued but shortages later when benefits are exhausted), breakfast routines could be disrupted.

Other routines could also be affected with consequential effects for children and youths. Meal routines seem especially important. Devault (1991) has described the centrality of meals as an “organizer of family life” (p. 38). The event of a meal not only involves feeding but also provides opportunities to converse, instruct, monitor family members, give stability, demonstrate affection, and generally produce “family.” Research confirms many of the benefits. For instance, regular family mealtimes have been found to be associated with better dietary intakes and healthier eating habits for adolescents (Neumark-Sztainer 2006), the transmission of culture to children (Ochs & Shohet 2006), language development (Snow & Beals 2006), improved psychological well-being among children (Fiese et al. 2006), and lower rates of youth substance abuse and risky behaviors

(National Center on Addiction and Substance Abuse 2010). Regular family routines, including eating meals together, are also associated with better control of children's asthma (Schreier & Chen 2010) and diabetes (Greening et al. 2007).

For this study, we use data from the Three-City Study, a longitudinal survey of 2,458 low-income families, to examine how family routines (including shared meals) and time use vary with participation in public and private food assistance programs. Although family routines have been studied extensively in other contexts, our study is one of only a handful that examine these processes and food assistance.

One study by Roy et al. (2012) examined how parents' and teenagers' time use differed with participation in different combinations of food assistance programs. Consistent with expectations, they found that parents spent less time preparing food and engaging in primary eating activities if their children received school meals but the family did not participate in the SNAP. However, they unexpectedly found that parents who participated in SNAP spent less time in primary eating activities and that adolescents who simultaneously participated in some but not all food assistance programs spent less time in primary eating activities. While provocative, a shortcoming of this research is that it used individual time-diary data from the American Time Use Survey (ATUS) that were limited to people aged 15 and over (e.g., excluded younger children and teenagers) and that had very limited information on interactions between parents and children. Additionally, the ATUS has relatively few demographic and household measures.

Another recent study by Waehrer (2008) used the Panel Study of Income Dynamics and found that children who received school meals reported fewer (primary)

eating activities, especially on weekdays. A shortcoming of Waehrer's study is that primary eating activities tend to be under-reported, especially when people are at school or at work. These reporting anomalies could explain the unexpected results. In a different context, a qualitative investigation of school breakfast clubs in the U.K. by Shemilt et al. (2003) found that parents in participating households reported fewer problems getting their children to eat regular breakfasts and reduced time pressure and stress on school mornings than other parents.

Our study uses a richer set of outcome measures than these earlier studies. Some of the outcome measures from the Three-City Study are based on one-day recall diaries describing the activities of both a parent and a child, while others are based on short recall and "usual-activity" questions. Another advantage of our study is that it examines measures that were collected longitudinally and prospectively. The availability of longitudinal data allows us to estimate fixed-effects specifications that can mitigate selection biases associated with omitted variables.

Conceptual Approach

Our empirical analysis is grounded in a household production framework (Becker 1965, Gronau 1977) that is modified to incorporate temporal routine (Hamermesh 2005). We begin with a standard household production model in which parents are assumed to have preferences over "commodities," such as meal consumption, household health outcomes, and their children's developmental outcomes, that are produced using inputs of time and goods. The activities that we examine can be viewed as time inputs, while food assistance from federal and private programs would be goods inputs. Households also

face constraints on their available time and budgets. Households choose activities and goods purchases to advance their preferences subject to the production, time, and budget constraints.

Food assistance programs (and transfer programs generally) are intended to improve well-being among low-income households by providing them with additional goods which should allow them to achieve better nutritional outcomes. However, Vickery (1977) has cautioned that it is important to consider goods needs, time needs, and their interrelationship when assessing poverty relief. Depending on the household production functions, the specific characteristics of the assistance, household preferences, and other characteristics, different types of food assistance could substitute for or complement certain types of household activities.

SNAP and WIC benefits can only be used for unprepared food and ingredients, many of which require additional preparation to form consumable meals. Thus, to take advantage of SNAP and WIC assistance, families must not only allocate time for meal consumption but also for meal preparation. Rose (2007) reviewed lunch and dinner recipes that were recommended as part of the Thrifty Food Plan (TFP, the basis for SNAP benefit guarantee) and estimated that preparing the recipes would require at least 16 hours a week; Rose's estimate likely understated the total time inputs of the TFP because it excluded shopping, clean-up, and the preparation of other non-recipe meals. Rose also examined time-use data and found that non-working women in SNAP households reported spending six more hours a week preparing meals than non-working women generally.

Changes in meal preparation could, in turn, affect other activities. On the one hand, more time in meal preparation could increase the availability of parents for secondary activities (activities that can be performed jointly with meal preparation). For instance, children could be playing quietly or doing their homework while caregivers are preparing meals. Through an effect on preparation times, the receipt of food assistance could increase these routines. On the other hand, increased preparation time associated with food assistance could cause families to substitute away from (could crowd out) other family activities in order to satisfy these new demands on their time. An obvious activity to consider is meal consumption. Analyses of meal times often fail to distinguish between preparation and consumption, but Woodward (2012) has shown that these activities are very different, with preparation having properties that we commonly associate with household production tasks and meal consumption having properties we associate with leisure activities. Consistent with this distinction, ethnographic evidence from a subset of the families in the Three-City Study indicates that some poor working mothers use meals from fast food restaurants as an explicit strategy to increase time available for consumption by reducing time needed for preparation (Tubbs et al. 2005). Other activities could also be affected by changes in meal preparation time.

In contrast to the benefits from the SNAP and WIC, assistance from the SBP and NSLP consists of complete meals, which do not involve time parental inputs and are also consumed outside the home. If we just consider breakfasts and lunches as household commodities, the SBP and NSLP would add to household resources, while reducing time pressures on parents. However, if family meal consumption (or preparation) is an activity

that is valued by itself or that contributes to the production of other commodities, the provision of school breakfasts and lunches could create pressures for parents to compensate with other activities, such as increasing the frequency with which they eat dinner together or engage in non-meal activities.

So far, we have considered the implications of food assistance in the context of the standard household production approach, which considers the incidence and amount of activities performed but not their timing. Hamermesh (2005) extended the standard model by assuming that (a) people have preferences regarding commodities and the timing of activities and (b) the production of commodities is made more or less costly by certain timing patterns. For our purposes, we could view “routine”—the same activity being repeated across days—as something that parents value but also as something that parents find costly to maintain.⁴¹ Some implications from this model are that more resources, either in terms of money or assistance, should also allow families to obtain more routine and that family activities will be strongly affected by external temporal cues, such as work or school schedules. Consistent with this, Crouter and McHale (1993) found that parental involvement and monitoring were associated with seasonal school schedules and parental employment. However, in another article that uses data from the Three-City Study, Coley et al. (2007) found that family routines were not associated with work or welfare transitions.

Within this framework, the possible effects of food assistance on household behaviors depend on parents’ valuations of routine and structure as opposed to variety.

⁴¹ Hamermesh actually examined the opposite case with his model, positing that people desired variety but found variety costly to produce.

Those valuations could differ with culture or class. For example, Lareau (2003) has described how less-affluent and working-class parents often rely on the “accomplishment of natural growth” as a child development strategy, which leads them to provide less supervision and encourage fewer structured activities for their children. Consistent with this, Kalenkoski et al. (2011) found that teenagers in disadvantaged circumstances spent less time in non-classroom educational activities but more time in unsupervised other activities. A perception that disadvantaged families lack or undervalue structure has motivated some policy analysts to recommend paternalistic assistance policies that mandate structure (see, e.g., Kane 1987 and Mead 1997). In this context, it is possible that participation in food assistance programs, like the School Breakfast Program, provide an external temporal cue that allows low-income households, like the ones in our sample, to add structure to their days and strengthen their family routines.

However, the costs of coordination could also differ across families. Roy et al. (2004) documented how break-downs in child care arrangements, irregular and non-standard work hours, a reliance on public transportation for work and shopping, and numerous institutional appointments interfered with poor families’ ability to structure their time. In this case, participation in food assistance program may have no effect on family routines because these families might not have the flexibility to adjust the use of their time and routines.

As this discussion indicates, standard theories lead to equivocal predictions. Possible effects of food assistance on family routines could vary depending on whether the assistance increases or decreases meal preparation activities, whether preparation

activities can be conducted jointly with other activities, and whether households place higher valuations on routine or variety. Previous research is also equivocal. Our empirical analysis will address the open question regarding the relationship between food assistance and family routines using a low-income sample.

Data

For our empirical analyses of family routines and time use we use data from the Three-City Study, a longitudinal survey of 2,458 children and their caregivers who were initially living in low-income neighborhoods in Boston, Chicago, and San Antonio. At the time of the first interview in 1999, the families all had incomes below 200 percent of the poverty line and had at least one child either 0-4 or 10-14 years old, whom the study designated as a “focal child.” Although the survey included many public assistance recipients, it was not specifically restricted to these groups. It was also designed to include poor and near-poor families, so there are many families who could potentially become eligible for assistance if their economic circumstances shifted. Thus, the study is especially well-suited for comparing households with and without different types of food assistance.

After the initial interviews, follow-up interviews were conducted in 2000-1 and 2005. Retention rates were high with 88 percent of the original sample participating in the second round and 80 percent participating in the third round. In each wave, interviews were conducted with both the focal child and the child’s caregiver. In cases where the child and caregiver separated, both were subsequently followed and interviewed.

For our analyses, we focus on households with focal children from the older cohort (10-14 years old at the initial interview). These children were school-aged in waves 1 and 2 and therefore potentially eligible to participate in school meal programs. The observations for these households could also be used in longitudinal analyses. In contrast, children from the younger cohort were not school-aged until wave 3 and could not be used in longitudinal analyses. We further focus on information provided by caregivers (almost always the focal children's mothers), who were asked questions regarding family routines, their children's schedules, household food assistance, and other characteristics of their households. We restrict our analyses to caregivers who were respondents in each of the first two waves of the Three-City Study, who continued to co-reside with the focal children (continuing caregivers), whose focal children were attending elementary school, middle school, junior school, high school or vocational school in the two waves, and who were interviewed during the school year (September-May) in both waves. We also drop observations with missing information for the outcome or explanatory variables. We exclude information from the third wave of the Three-City Study because many of the older cohort of focal children had already completed or left secondary school by this time and because these children were relatively old (generally 16 or older). We consider continuing respondents to facilitate longitudinal analyses, and we focus on households with enrolled children during the school year to examine the roles of school meal programs.

There were 1,158 households with focal children from the older cohort in the initial wave of the Three-City Study. Of those, 1,011 had continuing caregivers who were

also respondents in the second wave. Among those, there were 956 who had focal children attending elementary school, middle school, junior school, high school or vocational school in the first and second waves of the survey, and 627 were interviewed during the academic years in those waves. There were 521 caregivers with no item non-response to the questions used as explanatory variables, leaving a sample of 1,042 person-year observations. For our analyses of daily schedules, we further restrict the sample to reports that describe a weekday. Of the 521 caregivers that satisfy our baseline set of restrictions, 319 had reports describing a weekday on the first and second waves of the Three-City Study, resulting in 638 potential person-year observations.

Outcome measures. In each wave of the Three-City Study, caregivers were asked five questions regarding regular family routines. Two of the questions asked about meals: how often the family ate dinner/supper together and how often some of the family ate breakfast together in the morning. The possible responses included “almost never,” “sometimes,” “usually,” and “always.” In addition to meal routines, caregivers were also asked how often

- the family spent time talking or playing quietly,
- the children spent time doing homework, and
- the family observed consistent bedtimes for children.

Our analysis separately examines the responses to each of the questions regarding the frequency with which families engaged in particular types of routines. It also examines a composite, recoded measure that is formed by averaging the categorical responses from the five questions.

Caregivers were also asked about the periods of time that they and the focal children spent apart during the previous day. The survey asked about up to five spells apart. For each spell, caregivers provided the start and stop times, their locations, and their children's locations. We use these measures to examine the total time the caregiver spent apart from the focal child and the timing of caregiver's first separation from the child, if a separation occurred.

Food assistance. In each wave, the Three-City Study asked caregivers about their household's participation in several specific food assistance programs, including the SNAP, NSLP, and SBP, and WIC. In addition, caregivers were asked about their use of local food pantries and food banks. We create binary indicators for the receipt of these five types of assistance. These measures of government and private food assistance are the primary independent variables in the empirical analyses.⁴²

Other explanatory measures. Our analyses include a rich set of additional explanatory measures that are available from the Three-City Study and that may be correlated with the outcomes of interest and with food assistance. To account for economic resources, we include a measure of the household's total monthly "cash" income at the time of the interview from all sources. We measure welfare participation with an indicator for current receipt of benefits from the Temporary Assistance for Needy Families (TANF) program. To account for wealth, we include indicators for whether the household owned a home, vehicle, held financial assets, or had outstanding loans. We

⁴² Although the focal children in our analysis sample were too old to be eligible for WIC, they may have had other household members who were eligible. Woodward and Ribar (2012) report evidence that WIC assistance may be shared among household members.

measure the demographic composition of the household with variables for the number of adults in the household, the number of children, and the age of the youngest child.

We also include measures of the caregiver's work status. All models include indicators for whether the caregiver was working full time or part time. The health status of the caregiver is measured by an indicator for work- or activity-limiting disabilities and by a general self-reported index of health. Other characteristics of the caregiver such as marriage and cohabitation status, age, race and ethnicity, nativity, and education are also included. We also incorporate measures of social networks. We include binary measures of the extent of social networks that indicate whether the caregiver reports having enough people who will listen, provide childcare, help with small favors and loan money. Additionally, we include variables that control for the focal child's age and gender. Finally, there are indicators for the wave of the interview, the month of the interview, and the city of residence.

The Three-City Study used a stratified sampling design. In addition, the study experienced modest levels of attrition over time. To account for these design and attrition issues, all of our empirical analyses incorporate sampling weights that are provided with the study.

Descriptive Analysis

We examine unconditional means of the families' and children's outcomes and characteristics as well as means conditional on the receipt of different forms of food assistance. The means are reported in Table 3.1.

Comparisons of means reveal only small differences in family routines and daily schedules among households with different patterns of school meal participation. None of these differences is statistically significant. Differences do appear when we consider WIC receipt. WIC households have dinner together more often, have more consistent homework schedules, and observe more consistent bed times than non-WIC households. Also, children in WIC households separate from their caregivers about an hour earlier than those from non-WIC households. SNAP participation is associated with less frequent occurrences of the family spending time talking or playing quietly together.

The children differ in other characteristics. Children who receive school meals are more likely to live in households that receive TANF, have lower incomes, have fewer assets and financial accounts, and have more children. The caregivers of children that receive school meals are less likely to have a college education and less likely to be married with a spouse present. Families of children who receive school meals have more and younger children. Even though there are no citizenship requirements to participate in the school meal programs, the caregivers of children who receive them are less likely to be foreign born. Families that receive WIC have more and much younger children; caregivers are more likely to be younger than non-WIC households. Households that receive SNAP are more likely to receive welfare, have lower incomes, fewer assets and financial accounts, fewer adults, more and younger children than non-SNAP households. Also, caregivers in SNAP households are less likely to be high school graduates and less likely to be married with spouse present. Since 1997 people who are foreign born have faced restrictions on food stamp use; this is reflected in the data that shows that

caregivers who receive SNAP are less likely to be foreign-born. Because of the differences in the children's household, caregiver, and own characteristics, we also undertake multivariate analyses.

Multivariate Analyses

We estimate multivariate models of the family routine and time use outcomes that account for other observable characteristics and, where possible, some types of unobservable characteristics. For the multivariate analyses, let Y_{it} denote a measure of either a family routine or time use outcome for household i in survey wave t . Let \mathbf{F}_{it} be a vector of public and private food assistance measures for the same household and time period, and let \mathbf{X}_{it} be a vector of other observed characteristics of the household. We estimate two principal types of models.

First, we estimate regression models of the form

$$Y_{it} = \mathbf{A}_{OLS}' \mathbf{F}_{it} + \mathbf{B}_{OLS}' \mathbf{X}_{it} + \varepsilon_{it} \quad (3.1)$$

where ε_{it} represents unobserved characteristics of the household and \mathbf{A}_{OLS} and \mathbf{B}_{OLS} are vectors of coefficients to be estimated. In the model, the estimates of \mathbf{A}_{OLS} represent the conditional associations of food assistance with the household outcome. The specifications are pooled regression models.⁴³

Second, we also utilize the longitudinal data to estimate household fixed effect models. Suppose that the unobserved characteristics in ε_{it} can be decomposed into a time-

⁴³ For the family routines measures, which are ordered categorical outcomes, we have also estimated appropriate qualitative dependent variable models with results that are similar to those reported here.

invariant, household-specific component, η_i , and a time-varying component, v_{it} , such that $\varepsilon_{it} = \eta_i + v_{it}$. The fixed effects model can be written

$$Y_{it} = \mathbf{A}_{FE}' \mathbf{F}_{it} + \mathbf{B}_{FE}' \mathbf{X}_{it} + \eta_i + v_{it}. \quad (3.2)$$

The fixed effects procedure uses differencing to condition out the η_i . The resulting estimates of \mathbf{A}_{FE} represent associations of the food assistance programs with the outcome measures that condition on the observed characteristics and the household-specific unobserved characteristics, thus removing a potential source of biasing heterogeneity.

Although the fixed-effects procedure addresses some forms of omitted-variable bias, it does not address biases caused by time-varying unobserved characteristics that may be associated with family routines and the receipt of food assistance programs. For example, unmeasured changes in household needs that cause families to take up food assistance and change their family routines could still contribute biases.

The fixed-effects procedure also places demands on the data. Depending on the outcome under analysis there are between 256 and 503 households that satisfy our sample restrictions. Within these households, 15% of children switched their participation status in the SBP; 13% switched their participation in the NSLP; 16% switched their WIC participation, and 19% of families changed their SNAP participation. The modest sample sizes and limited number of program changes reduce the statistical power of the analyses and our ability to draw statistical inferences.

Ordinary Least Squares Results

Selected coefficients and standard errors from OLS models of family routines and time schedules are reported in Table 3.2. Specifically, Table 3.2 lists coefficients for participation in the public and private food assistance programs, participation in TANF, and total household cash income. Full results are reported in Appendix G.

As with the descriptive analysis, participation in the SBP is only weakly associated with family routines and daily schedules. None of the associations is statistically significant. Several of the coefficients for participation in the NSLP are larger in magnitude. For instance, NSLP participation is negatively associated with breakfast routines, though it falls just short of being statistically significant. NSLP participation is also negatively associated with dinner routines and positively associated with total hours apart, albeit with imprecise estimates that are consistent with positive or negative effects.

Also consistent with the descriptive results, SNAP participation is negatively associated with family-time routines, with estimates that can be statistically distinguished from zero. SNAP participation is also associated with significantly later daily separation times for children. WIC participation is significantly associated with fewer breakfast routines, more dinner routines, more homework routines, and more family-time routines.

Food pantry use is not significantly associated with any of the family routines in the OLS models. However, it is associated with earlier daily separation times. Similarly, TANF participation is not significantly associated with family routines or daily schedules, which is similar to the findings of Coley et al. (2007). Higher levels of household income are associated with fewer dinner routines, less consistent bed times, and fewer family

routines generally. These results are consistent with Hamermesh's (2005) finding that variety is a normal good.

Longitudinal Fixed-Effect Results

A concern with the preceding estimates is that unobserved characteristics that are related to the outcome measures and food assistance could be confounding the results. If, for instance, families with good organization skills are more likely to receive food assistance and also more likely to have strong family routines, then OLS coefficient estimates could be over-estimated. On the other hand, if families with good organization skills are less likely to receive food assistance program, then OLS coefficient estimates could be under-estimated. Because of the availability of longitudinal data, we can address potential biases associated with time-invariant characteristics by estimating fixed effects models. Estimated coefficients and standard errors from those models are reported in Table 3.3.

Estimates from the fixed effects models indicate that participation in the SBP is not strongly associated with family routines. However, SBP participation is associated with earlier separation times for children, a result that is consistent with mothers leaving home earlier to work or children leaving home earlier to take advantage of school breakfasts. The size of the effect seems more consistent with the work time explanation. SBP participation is also associated with longer times apart, though the estimate is imprecise. Participation in the school lunch program is also only modestly associated with family routines and more strongly associated with earlier and longer separation times; all of the associations, however, are insignificant.

SNAP participation is significantly negatively associated with homework routines. There is also evidence that SNAP participation may be negatively associated with dinner routines. WIC participation is significantly positively associated with dinner routines, homework routines, family-time routines, and family routines generally. Food pantry use is associated with fewer family-time routines and earlier separation times.

TANF receipt is only weakly associated with family routines but is strongly associated with later separation times and fewer hours apart. In contrast, higher levels of family income are associated with fewer dinner routines, fewer family routines overall, and more hours apart.

Among the other characteristics (Appendix H), the wealth measures are negatively correlated with a number of family routines. More minors in the household are associated with stronger family routines, a greater frequency of eating breakfast together, more homework routines, and earlier separation times. The age of the youngest child is positively associated with stronger family routines, more dinner routines, more bedtime routines, and earlier separation times. Full-time work by the caregiver is associated with a greater frequency of eating breakfast together, but with less frequent homework routines. Part-time work is also associated with less frequent homework routines, and with later separation times, which might be consistent with evening work. Caregivers who are married with a spouse present spent less time separated from the focal child and have later separation times. Cohabiting caregivers report more homework and bedtime routines than single caregivers.

Sensitivity Analyses

One potential concern is that participation in multiple programs at the same time creates multicollinearity problems preventing to find statistically significant associations between food assistance programs and the outcomes described in this paper. Eligibility requirements are the same for the SBP and NLSP, and children living in families that receive SNAP are also eligible to receive reduced price meals at school. However, the Variance Inflation Factor for each one of the food assistance variables is less than 2 suggesting that our results are not affected by multicollinearity.

Discussion

Despite evidence linking family routines to positive health and developmental outcomes for children, such routines have not been examined in studies of the broader impacts of food assistance programs in the U.S. Data from the Three-City study allow us to address that gap. In longitudinal fixed effects analyses, we find statistically significant evidence that participation in the SNAP is associated with fewer homework routines and less precise evidence that participation in SNAP is associated with fewer dinner routines. These negative correlations were also observed in linear regression models, although the coefficient estimates in those models were smaller and not precisely estimated. The negative effect on these routines could be a consequence of an increase in meal preparation time, which causes caregivers to substitute away from dinner consumption and time spent in homework activities.

The longitudinal models indicate that WIC participation is associated with the family eating dinner together more frequently, more homework routines, higher

frequency of the family talking or playing quietly together, and more family routines in general. Many of these positive associations were also observed in the linear regressions models and could be the result of a compensating effect. WIC provides food ingredients for the younger children in the household, so it is possible that caregivers compensate the differential in food availability by spending more time with members of the household that are not direct beneficiaries of the WIC assistance. The positive association between WIC and dinner routines could also be the result of the nutritional education component of the program. One of the main objectives of this component is to “assist individuals in nutritional risk achieve a positive change in dietary habits” (U.S. Food and Nutrition Service, 2010).

Both linear regression models and longitudinal models suggest only a weak association between school meals and family routines. However, longitudinal models suggest that children who participate in the School Breakfast Program separate earlier from their caregivers.

Food pantry use is estimated to be negatively associated with time spent talking or quietly playing with children and with earlier separation times. The availability of pantry may be less convenient and its use could interfere with family routines.

Many of the associations between food assistance programs, family routines and time use described in the paper are observed both in the linear regression and longitudinal models. However, some cautions need to be applied to our analyses. As we discuss, the longitudinal models do not address the bias caused by time-varying unobservable characteristics correlated with food assistance program participation and family routines.

Our analyses also only consider relatively older children, who may not have family routines as strong as the ones of younger children. Finally, the analysis in this paper could also benefit of larger sample sizes with larger variation in the independent variables of interest, which would allow to identify effects with more precision.

Tables and Figures

Table 3.1 – Means of Outcomes and Independent Variables

	All	0 School meals	1 School meal	2 School meals	No WIC	WIC	No SNAP	SNAP
Outcomes ^A								
Family routines index	2.748	2.849	2.662	2.752	2.726	2.885	2.782	2.696
Breakfast routines	2.254	2.571	2.546	2.195	2.272	2.146	2.227	2.295
Dinner routines	2.885	2.972	2.637	2.909	2.825	3.246***	2.947	2.788
Homework routines	3.005	3.149	3.011	2.994	2.970	3.220*	3.054	2.929
Bedtime routines	3.022	3.009	2.838	3.046	2.978	3.292**	3.015	3.033
Family time routines	2.833	2.844	2.627	2.857	2.829	2.856	2.940	2.666**
Total hours apart	7.917	7.996	7.910	7.914	7.773	8.525	8.238	7.459
Start time first separation	8.763	8.490	9.509	8.701	8.976	7.987*	8.439	9.212
Household characteristics								
TANF participation	0.299	0.041	0.346***	0.311***	0.287	0.370	0.073	0.631***
Monthly income (\$1000s)	1.441	2.098	1.285***	1.415***	1.422	1.550	1.623	1.175***
Own home	0.236	0.447	0.265	0.217**	0.238	0.222	0.343	0.079***
Own car	0.422	0.716	0.265***	0.423***	0.417	0.453	0.572	0.203***
Has financial accounts	0.359	0.633	0.239***	0.355***	0.383	0.218**	0.468	0.199***
Has outstanding loans	0.461	0.608	0.401	0.458	0.479	0.354	0.550	0.331***
Number of adults	1.856	2.507	1.744***	1.825***	1.877	1.733	2.035	1.594***
Number of minors	2.857	2.039	2.599	2.949***	2.659	4.020***	2.563	3.288***
Age of the youngest child	8.1	9.8	8.6	7.9*	8.9	3.3***	8.4	7.6*
Boston	0.108	0.267	0.104**	0.098**	0.109	0.105	0.124	0.086**
Chicago	0.649	0.423	0.809***	0.644**	0.669	0.536	0.565	0.773***

Wave 2	0.500	0.672	0.396**	0.502*	0.506	0.465	0.513	0.481
Caregiver characteristics								
Works full time	0.356	0.499	0.305	0.352	0.346	0.409	0.442	0.228***
Works part time	0.169	0.255	0.082	0.174	0.179	0.108	0.170	0.167
High school graduate	0.370	0.166	0.308	0.393***	0.405	0.167***	0.414	0.307*
College education	0.191	0.504	0.323	0.151***	0.195	0.168	0.244	0.113***
Married spouse present	0.298	0.506	0.286	0.285**	0.284	0.382	0.388	0.168***
Cohabiting	0.083	0.012	0.152*	0.079***	0.083	0.084	0.084	0.083
Foreign born	0.179	0.488	0.061***	0.173***	0.181	0.167	0.254	0.070***
Health status	2.881	2.798	3.086	2.859	2.918	2.660	2.747	3.077***
Disability status	0.267	0.208	0.264	0.271	0.293	0.114***	0.213	0.346**
Black	0.582	0.186	0.714***	0.592***	0.595	0.508	0.462	0.758***
Hispanic	0.389	0.731	0.243***	0.385***	0.374	0.479	0.504	0.221***
Age	39.4	42.2	42.0	38.8***	40.3	33.8***	39.4	39.3
People who will listen	0.488	0.618	0.404	0.491	0.492	0.470	0.488	0.488
People who will help with childcare	0.494	0.576	0.469	0.492	0.498	0.468	0.493	0.495
People who helps with small favors	0.480	0.595	0.405	0.482	0.465	0.568	0.463	0.506
People who will loan money	0.386	0.485	0.376	0.380	0.394	0.341	0.417	0.341
Child characteristics								
Female	0.549	0.799	0.564**	0.529***	0.544	0.573	0.515	0.598
Age	12.8	13.4	12.8	12.7***	12.8	12.8	12.9	12.7
Proportion of observations	1.000	0.059	0.111	0.831	0.854	0.146	0.594	0.406

Note: Authors' calculations based on 1,042 person-year observations of school-enrolled children living with a continuing caregiver interviewed the first and second waves of the Three-City Study. All of the statistics were calculated using sample weights. Asterisks indicate whether the means are significantly different for children receiving food assistance relative to those who do not. *** $p < 0.01$, ** $p < 0.05$, $p < 0.1$.

^A The sample sizes used to calculate mean outcomes are smaller than the one used to calculate mean controls.

Table 3.2 – Selected Estimates from Ordinary Least Squares Models of Family Routines and Daily Schedules

VARIABLES	Family routines	Breakfast routines	Dinner routines	Homework routines	Bedtime routines	Family time routines	Total hours apart	Start time first separation
School Breakfast Program	0.070 (0.120)	-0.119 (0.203)	0.146 (0.163)	0.068 (0.180)	0.004 (0.191)	0.250 (0.180)	-0.055 (1.557)	0.090 (0.693)
National School Lunch Program	-0.172 (0.158)	-0.448 (0.273)	-0.271 (0.190)	-0.262 (0.243)	0.154 (0.174)	-0.121 (0.222)	1.525 (1.347)	0.055 (1.009)
SNAP	-0.053 (0.088)	0.113 (0.142)	-0.186 (0.170)	-0.119 (0.121)	0.104 (0.148)	-0.243* (0.124)	-0.263 (1.023)	1.353** (0.684)
WIC	0.157 (0.098)	-0.311** (0.136)	0.492*** (0.172)	0.387** (0.151)	0.196 (0.171)	0.254* (0.144)	0.969 (0.966)	-0.081 (0.612)
Food pantry	-0.013 (0.143)	0.245 (0.210)	0.080 (0.196)	0.128 (0.203)	-0.159 (0.187)	-0.220 (0.165)	-0.123 (0.897)	-1.400* (0.728)
TANF participation	-0.040 (0.105)	-0.011 (0.174)	0.025 (0.178)	0.024 (0.152)	-0.136 (0.153)	-0.037 (0.143)	-0.503 (1.057)	-0.859 (0.802)
Monthly income (\$1000s)	-0.085** (0.035)	-0.079 (0.055)	-0.106** (0.052)	0.005 (0.064)	-0.077* (0.045)	-0.080 (0.059)	0.244 (0.362)	0.470 (0.321)
Number of observations	1,006	1,006	1,006	1,006	1,006	1,006	614	512
Number of families	503	503	503	503	503	503	307	256

Note: Estimates from authors' linear regressions using school-enrolled children living with a continuing caregiver interviewed in the first and second waves of the Three-City Study. Regressions also include controls for home ownership; car ownership; whether the household has financial assets and outstanding loans; number of adults; number of children; age of the youngest child; city of residency; survey wave; month of interview; caregiver's full time and part time work status, disability status, self-reported health, marriage status, cohabitation status, age, race, ethnicity, nativity, and education; access to people who will listen, provide childcare, help with small favors and loan money; and the focal child's age and gender. Estimates incorporate sampling weights. Robust standard errors are shown in parentheses.

***Coefficient different from zero $p < 0.01$ **Coefficient different from zero $p < 0.05$ *Coefficient different from zero $p < 0.1$

Table 3.3 – Selected Estimates from Fixed Effects Models of Family Routines and Daily Schedules

VARIABLES	Family routines	Breakfast routines	Dinner routines	Homework routines	Bedtime routines	Family time routines	Total hours apart	Start time first separation
School Breakfast Program	0.162 (0.162)	0.177 (0.201)	0.008 (0.219)	0.006 (0.248)	0.174 (0.275)	0.291 (0.300)	1.343 (1.741)	-2.381** (1.051)
National School Lunch Program	0.012 (0.165)	-0.255 (0.273)	-0.025 (0.207)	-0.005 (0.360)	0.125 (0.214)	0.201 (0.299)	1.734 (1.961)	-1.843 (1.518)
SNAP	-0.021 (0.134)	0.200 (0.190)	-0.323 (0.208)	-0.324* (0.166)	-0.032 (0.194)	0.069 (0.271)	-0.737 (1.336)	0.079 (0.864)
WIC	0.321* (0.166)	-0.248 (0.224)	0.588** (0.257)	0.773*** (0.251)	0.464 (0.324)	0.481** (0.205)	1.102 (1.211)	-1.196 (0.998)
Food pantry	-0.018 (0.138)	0.104 (0.212)	0.193 (0.262)	0.146 (0.217)	0.003 (0.174)	-0.372* (0.191)	0.068 (1.327)	-3.778*** (1.045)
TANF participation	-0.058 (0.165)	-0.033 (0.288)	-0.084 (0.229)	0.134 (0.238)	0.082 (0.189)	-0.197 (0.252)	-5.262*** (1.456)	2.087** (0.835)
Monthly income (\$1000s)	-0.080* (0.047)	-0.099 (0.068)	-0.117* (0.066)	0.002 (0.078)	-0.029 (0.068)	-0.074 (0.084)	1.227** (0.483)	-0.055 (0.396)
Number of observations	1,006	1,006	1,006	1,006	1,006	1,006	614	512
Number of families	503	503	503	503	503	503	307	256

Estimates from authors' fixed effects regressions using school-enrolled children living with a continuing caregiver interviewed in the first and second waves of the Three-City Study. Regressions also include controls for home ownership; car ownership; whether the household has financial assets and outstanding loans; number of adults; number of children; age of the youngest child; survey wave; month of interview; caregiver's full time and part time work status, disability status, self-reported health, marriage status, and cohabitation status; and access to people who will listen, provide childcare, help with small favors and loan money. Estimates incorporate sampling weights. Robust standard errors are shown in parentheses.

***Coefficient different from zero $p < 0.01$ **Coefficient different from zero $p < 0.05$ *Coefficient different from zero $p < 0.1$

REFERENCES

- Abadie A., Diamond A., and Hainmueller J. (2010). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association*, 105(490), 493-505.
- Abadie A., Diamond A., and Hainmueller J. (2011). SYNTH: Stata Module to Implement Synthetic Control Methods for Comparative Case Studies. Software item, available <http://econpapers.repec.org/software/bocbocode/s457334.htm>.
- Ai C. and Norton E. (2003). Interaction Terms in Logit and Probit Models. *Economics Letters*, 80(1), 123-129.
- Aizer, A. (2007). Public Health Insurance, Program Take-Up, and Child Health. *Review of Economics and Statistics*, 89(3), 400-415.
- Almond, D., and Currie, J. (2010). Human Capital Development Before Age Five. *NBER Working Paper*, 15827.
- Angel, R., Lein L., and Henrici, J. (2006). Poor families in America's Health Care Crisis, Cambridge University Press.
- Angrist J. and Pischke J. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.
- Arrow K. (1963). Uncertainty and the Demand for Medical Care. *American Economic Review*, 53(5), 941-73.
- Aughinbaugh, A. (2004). The Impact of Attrition on the Children of the NLSY79. *Journal of Human Resources*, 39(2), 536-563.

- Baker Institute. (2009). The Economic Impact of Uninsured Children on America, *James A. Baker III Institute for Public Policy of Rice University*, No. 40.
- Baum, C. (2003). Does Early Maternal Employment Harm Child Development? An Analysis of the Potential Benefits of Leave Taking. *Journal of Labor Economics* 21 (2), 409–448.
- Baum, C. (2011). “The Effects of Food Stamps on Obesity.” *Southern Economic Journal* 77: 623-51.
- Becker, G. (1965). “A Theory of the Allocation of Time.” *Economic Journal* 75: 493-517.
- Bertrand M., Duflo E., and Mullainathan S. (2004). How Much Should We Trust Difference-in-Differences Estimates? *Quarterly Journal of Economics*, 119(1), 249-275.
- Bhattacharya J., Bundorf M.K., Pace N., and Sood N. (2011). Does Health Insurance Make You Fat? In Grossman M. and Mocan N. (eds.), *Economic Aspects of Obesity*, Chicago, IL: University of Chicago Press, 35-64.
- Black D., M. Berger, and F. Scott. (2000). Bounding Parameter Estimates with Non-classical Measurement Error. *Journal of the American Statistical Association*, 95(451), 739-748.
- Blue Cross Blue Shield of Massachusetts (2006). Massachusetts Health Care Reform Bill Summary. Available at <http://tinyurl.com/bn5yepz>.
- Bound, J., and Solon, G. (1999). Double the Trouble: On the Values of Twins Based Estimation of the Return to Schooling. *Economics of Education Review*, 18, 169-182.
- Card D., Dobkin C., and Maestas N. (2009). Does Medicare Saves Lives? *Quarterly Journal of Economics*, 124(2), 597-636,
- Card D., Dobkin C., and Maestas N. (2004). The Impact of Nearly Universal Insurance Coverage on Health Care Utilization: Evidence from Medicare. *National Bureau of Economic Research Working Paper* #10365.

- Case, A., and Paxson, C. (2010). Causes and Consequences of Early-life Health. *Demography*, 47S, S65-S85.
- Case, A., Fertig, A., and Paxson, C. (2005). The Lasting Impact of Childhood Health and Circumstance. *Journal of Health Economics*, 24(2), 365-89.
- Cawley J. (2004). The Impact of Obesity on Wages. *Journal of Human Resources*, 39(2), 451-474.
- Chen, Y., and Zhe Jin G. (2012). Does Health Insurance Coverage Lead to Better Health and Educational Outcomes? Evidence from Rural China, *Journal of Health Economics*, 31(1), 1-14.
- Cogan J., Hubbard R., and Kessler D. (2010). The Effect of Massachusetts' Health Reform on Employer-Sponsored Insurance Premiums. *Forum for Health Economics and Policy*, 13(2), Article 5.
- Copper, R. L. Goldenberg, R. L., Creasy, R. K., Dubard, M. B., Davis, R. O., Entman, S. S., Iams, J. D. and Cliver, S. P. (1994). A Multicenter Study of Preterm Birth Weight and Gestational Age-specific Neonatal Mortality. *American Journal of Obstetrics and Gynecology*, 170, 960-961.
- Courtemanche C. (2009). Rising Cigarette Prices and Rising Obesity: Coincidence or Unintended Consequence? *Journal of Health Economics*, 28(4), 781-798.
- Courtemanche C., Heutel G., and McAlvanah P. (2011). Impatience, Incentives, and Obesity. *National Bureau of Economic Research Working Paper* #17483.
- Cunha, F., and Heckman, J. (2008). Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Journal of Human Resources*, 43(4), 738-782.
- Cunha, F., Heckman, J., and Schennach, S. (2010). Estimating the Technology of Cognitive and Noncognitive Skill Formation, *Econometrica*, 78, (3), 883-931.

- Currie, J. (2000). Do Children of Immigrants Make Differential Use of Public Health Insurance?. *Issues in the Economics of Immigration*, editor George J. Borjas, University of Chicago Press.
- Currie, J. (2009). Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development. *Journal of Economic Literature*, 47(1), 87-122.
- Currie, J., Decker, S., and Lin, W. (2008). Has Public Health Insurance for Older Children Reduced Disparities in Access to Care and Health Outcomes?. *Journal of Health Economics*, 27(6), 1567–1581.
- Currie, J. and Gruber, J. (1996a). Health Insurance Eligibility, Utilization of Medical Care, and Child Health. *Quarterly Journal of Economics*, 111(2), 431-466.
- Currie, J. and Gruber, J. (1996b). Saving Babies: The Efficacy and Cost of Recent Changes in the Medicaid Eligibility of Pregnant Women. *Journal of Political Economy*, 104(6), 1263–1296.
- Currie, J., and Gruber, J. (1997). The Technology of Birth: Health Insurance, Medical Interventions, and Infant Health. *NBER working paper*, 5985.
- Currie, J., Stabile, M., Manivong, P., and Roos, L. (2010). Child Health and Young Adult Outcomes. *The Journal of Human Resources*, 45(3), 517-548.
- Currie, J., and Thomas, D. (1995). Medical Care for Children: Public Insurance, Private Insurance, and Racial Differences in Utilization, *Journal of Human Resources*, 30(1), 135-162.
- Cutler, D. (2003). Employee Costs and the Decline in Health Insurance Coverage, *Forum for Health Economics and Policy*, 6(3), 27-53.
- Coley, R., B. Lohman, E. Votruba-Drzal, L. Pittman, P. Chase-Lansdale. (2007). “Maternal functioning, time, and money: The world of work and welfare”. *Children and Youth Services Review* 29: 721-41.

- Crouter, A.C., and S.M. McHale. (1993). "Temporal Rhythms in Family Life: Seasonal Variation in the Relation Between Parental Work and Family Processes." *Developmental Psychology* 29: 198-205.
- Dafny, L., and Gruber, Jonathan. (2000). Does Public Insurance Improve the Efficiency Of Medical Care? Medicaid Expansions And Child Hospitalizations. *NBER Working Paper*, 7555.
- Dafny L. and Gruber J. (2005). Public Insurance and Child Hospitalizations: Access and Efficiency Effects. *Journal of Public Economics*, 89, 109-129.
- Dave D. and Kaestner R. (2009). Health Insurance and Ex Ante Moral Hazard: Evidence from Medicare. *International Journal of Health Care Finance and Economics*, 9, 367-390.
- Decker, S. (2007). Medicaid Physician Fees and the Quality of Medical Care of Medicaid Patients in the USA. *Review of Economics of the Household*, 5(1), 95-112.
- DeSalvo, K. B., Bloser, N., Reynolds, K., He, J., and Muntner, P. (2006). Mortality prediction with a single general self-rated health question. A meta-analysis. *Journal of General Internal Medicine*, 21(3), 267–275.
- Devault, M. 1991. *Feeding the Family: The Social Organization of Caring as Gendered Work*. Chicago: University of Chicago Press.
- Doyle J. (2005). Health Insurance, Treatment and Outcomes: Using Auto Accidents as Health Shocks. *Review of Economics and Statistics*, 87(2), 256-270.
- Fiese, B.H., K.P. Foley, and M. Spagnola. (2006). "Routine and Ritual Elements in Family Mealtimes: Contexts for Child Well-Being and Family Identity." *New Directions for Child and Adolescent Development* 111: 67-89.
- Finkelstein A. and McKnight R. (2008). What Did Medicare Do? The Initial Impact of Medicare on Mortality and Out of Pocket Medical Spending. *Journal of Public Economics*, 92, 1644-1669.

- Finkelstein A., Taubman S., Wright B., Bernstein M., Gruber J., Newhouse J., Allen H., Baicker K., and The Oregon Health Study Group (2011). The Oregon Health Insurance Experiment: Evidence from the First Year. *Quarterly Journal of Economics*, forthcoming.
- Fitzpatrick M. (2008). Starting School at Four: The Effect of Universal Pre-Kindergarten on Children's Academic Achievement. *B.E. Journal of Economic Analysis and Policy*, 8(1), Article 46.
- Fox, M., W. Hamilton, and B. Lin. (2004). *Effects of Food Assistance and Nutrition Programs on Nutrition and Health: Vol. 3, Literature Review*. Food Assistance and Nutrition Research Report no. 19-3. Washington, DC: USDA Economic Research Service.
- Fryer, R., and S. Levitt. (2004). Understanding the Black-White Test Score Gap in the First Two Years of School. *Review of Economics and Statistics*, 86: 447-464.
- Griliches, Z. (1979). Sibling Models and Data in Economics: Beginnings of a Survey. *Journal of Political Economy*, 87(5), S37-64.
- Greening, L., L. Stoppelbein, C. Konishi, S. S. Jordan, and G. Moll. 2007. "Child Routines and Youths' Adherence to Treatment for Type 1 Diabetes." *Journal of Pediatric Psychology* 32: 437-47.
- Gronau, R. 1977. "Leisure, Home Production, and Work—the Theory of the Allocation of Time Revisited." *Journal of Political Economy* 29: 1099-123.
- Grossman, M. (1972). On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy*, 80(2), 223-255.
- Gruber, J. (2000). Medicaid. *National Bureau of Economic Research Working Paper* #7829.
- Gruber J. (2008a). Incremental Universalism for the United States: The States Move First? *Journal of Economic Perspectives*, 22(4), 51-68.

- Gruber J. (2008b). Massachusetts Health Care Reform: The View From One Year Out. *Risk Management and Insurance Review*, 11(1), 51-63.
- Gruber J. (2010). Health Care Reform is a “Three-Legged Stool”: The Costs of Partially Repealing the Affordable Care Act. Center for American Progress. Available http://www.americanprogress.org/issues/2010/08/pdf/repealing_reform.pdf.
- Gruber J. (2011a). The impact of the Affordable care Act: How reasonable are the projections? *National Bureau of Economic Research Working Paper* #17168.
- Gruber J. (2011b). Massachusetts Points the Way to Successful Health Care Reform. *Journal of Policy Analysis and Management*, 30(1), 184-192.
- Hadley J. (2003). Sicker and Poorer—The Consequences of Being Uninsured: A Review of the Research on the Relationship between Health Insurance, Medical Care Use, Health, Work, and Income. *Medical Care Research and Review*, 60(2), 3-75.
- Hamermesh, D.S. (2005). “Routine.” *European Economic Review* 49: 29-53.
- Harrington S. (2010). U.S. Health Care Reform: The Patient Protection and Affordable Act. *Journal of Risk and Insurance*, 77(3), 703-708.
- Haveman R., and B. Wolfe. (1995). The Determinants of Children's Attainments: A Review of Methods and Findings, *Journal of Economic Literature*, 33(4), 1829-78.
- Heckman, J, and B. Singer. (1984). “A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data.” *Econometrica* 52: 271-320.
- Heckman J., J. Stixrud and S. Urzua. (2006). The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, 24(3), 411-482.
- Hoynes H., and E. Luttmer. (2011). The Insurance Value of State Tax-and-Transfer Programs. *Journal of Public Economics*, 95(11): 1466-1484.

- Hwang, A., Hwang, M., Xie H., Hardy B., and Skaggs D. (2005), Access to Urologic Care for Children in California: Medicaid versus Private Insurance. *Urology*, 66, 170–173.
- Idler E. and Benyamini Y. (1997). Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies. *Journal of Health and Social Behavior*, 38 (1), 21-37.
- Johnson-Lans S. (2005). *A Health Economics Primer*. Boston, MA: Addison Wesley/Pearson.
- Kalenkoski, C.M., D.C. Ribar, and L.S. Stratton. (2011). “How Do Adolescents Spell Time Use? An Alternative Methodological Approach for Analyzing Time Diary Data.” *Research in Labor Economics* 33: 1-44.
- Kane, T.J. (1987). “Giving Back Control: Long-Term Poverty and Motivation.” *Social Service Review* 61: 405-19.
- Kenney G., Cook A., and Dubay L. (2009). Progress Enrolling Children in Medicaid/CHIP: Who is Left and What are the Prospects for Covering More Children?. Timely Analysis of Immediate Health Policy Issues. Urban Institute. Retrieved from <http://www.urban.org/publications/411981.html>
- Kirk A. (2000). Riding the Bull: Experience with Individual Market Reform in Washington, Kentucky, and Massachusetts. *Journal of Health Politics, Policy and Law*, 25(1), 133-173.
- Kolstad J. and Kowalski A. (2010). The Impact of Health Care Reform on Hospital and Preventive Care: Evidence from Massachusetts. National Bureau of Economic Research Working Paper #16012.
- Kolstad J. and Kowalski A. (2012). Mandate-Based Health Reform and the Labor Market: Evidence from the Massachusetts Reform. *National Bureau of Economic Research Working Paper #17933*.
- Lareau, A. 2003. *Unequal Childhoods: Class, Race, and Family Life*. Berkeley, CA: University of California Press.

- Levine, P., and Schanzenbach, D. (2009). The Impact of Children's Public Health Insurance Expansions on Educational Outcomes. *Frontiers in Health Policy Research*, 12(1).
- Levy H. and Meltzer D. (2008). The Impact of Health Insurance on Health. *Annual Review of Public Health*, 29, 399–409.
- Long, S. (2008). On the Road to Universal Coverage: Impacts of Reform in Massachusetts at One Year. *Health Affairs*, 27(4), w270-84.
- Long, S. and Stockely K. (2011). The Impacts of State Health Reform Initiatives on Adults in New York and Massachusetts. *Health Services Research*, 46(1,II), 365-428.
- Long S., Stockely K. and Yemane A. (2009). Another Look at the Impacts of Health Reform in Massachusetts: Evidence Using New Data and a Stronger Model. *American Economic Review Papers and Proceedings*, 99(2), 508-511.
- Manning W., Newhouse J., Duan N., Keeler E., Leibowitz A., and Marquis M. (1987). Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment. *American Economic Review*, 77(3), 251-77.
- Massachusetts Division of Health Care, Finance, and Policy (2010). Health Care in Massachusetts: Key Indicators. Available <http://archives.lib.state.ma.us/handle/2452/69933>.
- Massachusetts Governor's Budget FY2010 (2009). Health Care. Retrieved from <http://www.mass.gov/bb/h1/fy10h1/prnt10/exec10/pbuddevhc.htm>.
- Massachusetts Health Insurance Connector Authority (2008). Report to the Massachusetts Legislature Implementation of the Health Care Reform Law, Chapter 58 2006-2008. Available <http://tinyurl.com/74osvco>.
- McDonough J., Rosman B., Phelps F., and Shannon M. (2006) The Third Wave of Massachusetts Health Care Access Reform. *Health Affairs*, 25(6), w420-31.

- Mead, L.M. (ed.). 1997. *The New Paternalism: Supervisory Approaches to Poverty*. Washington: Brookings Institution Press.
- Merton, M., A. Williams, and L. Shriver. (2009). "Breakfast Consumption in Adolescence and Young Adulthood: Parental Presence, Community Context, and Obesity." *Journal of the American Dietetic Association* 109: 1384-91.
- Meyerhoefer, C., and Y. Pylypchuk. (2008) "Does Participation in the Food Stamp Program Increase the Prevalence of Obesity and Health Care Spending?" *American Journal of Agricultural Economics* 90: 287-305.
- Miller S. (2011a). The Effect of Insurance on Outpatient Emergency Room Visits: An Analysis of the 2006 Massachusetts Health Reform. Working paper, University of Illinois.
- Miller S. (2011b). The Impact of the Massachusetts Health Care Reform on Health Care Use Among Children. Forthcoming, *American Economic Review Papers and Proceedings*.
- National Center on Addiction and Substance Abuse. (2010). *The Importance of Family Dinners IV*. New York: Columbia University.
- Neumark-Sztainer, D. (2006). "Eating Among Teens: Do Family Mealtimes Make a Difference for Adolescents' Nutrition?" *New Directions for Child and Adolescent Development* 111: 91-105.
- Nord, M., A. Coleman-Jensen, M. Andrews, and S. Carlson. (2010). *Household Food Security in the United States, 2009*. Economic Research Report No. (ERR-108). Washington, DC: USDA Economic Research Service.
- Nord, M., and A. Golla. (2009). *Does SNAP Decrease Food Insecurity? Untangling the Self-Selection Effect*. Economic Research Report No. (ERR-85). Washington, DC: USDA Economic Research Service.
- Ochs, E., and M. Shohet. (2006). "The Cultural Structuring of Mealtime Socialization." *New Directions for Child and Adolescent Development* 111: 35-49.

- Orzechowski W., Walker R. (2010). *The Tax Burden on Tobacco: Historical Compilation, Volume 45*. Arlington, VA: Orzechowski and Walker.
- Patel K., McDonough J. (2010). From Massachusetts to 1600 Pennsylvania Avenue: Aboard the Health Reform Express. *Health Affairs*, 29(6), 1106-1110.
- Pauly M. (1968). The Economics of Moral Hazard: Comment. *American Economic Review*, 58(3), 531-36.
- Pauly M. (2005). Effects of Insurance Coverage on Use of Care and Health Outcomes for Non-poor Young Women. *American Economic Review, Papers and Proceedings*, 95(2), 219-223.
- Peterson, K., M.L. Davison, K. Wahlstrom, J. Himes, M. Stevens, Y.S. Seo, M.L. Irish, K. Holleque, J. Haring, A. Hansen. (2004). "Fastbreak to Learning School Breakfast Program: A Report of the Fourth Year Results, 2002-03." University of Minnesota.
- Phillips A., Der G., and Carroll D. (2010). Self-reported health, self-reported fitness, and all-cause mortality: Prospective cohort study. *British Journal of Health Psychology*, 15(2), 337-346.
- Puhani P. (2008). The Treatment Effect, the Cross Difference, and the Interaction Term in Nonlinear "Difference-in-Differences" Models. IZA Discussion Paper Series, 3478.
- Raymond A. (2009). Massachusetts Health Reform: The Myth of Uncontrolled Costs. Massachusetts Taxpayer Foundation. Retrieved from <http://www.masstaxpayers.org/files/Health%20care-NT.pdf>
- Read, J. S., Troendle, J. F. and Klebanoff, M. A. (1997). Infectious Disease Mortality Among Infants in the United States, 1983 through 1987. *American Journal of Public Health*, 87, 192-198.
- Ribar, D., and L. Haldeman. (2011). "Universal-Free and Eligibility-Based School Breakfast Programs in Guilford County, North Carolina: Student Outcomes." Unpublished manuscript, University of North Carolina at Greensboro.

- Rose, D. (2007). "Food Stamps, the Thrifty Food Plan, and Meal Preparation: The Importance of the Time Dimension for US Nutrition Policy." *Journal of Nutrition Education and Behavior* 39: 226-32.
- Roy, K.M., C.Y. Tubbs, and L.M. Burton. (2004). "Don't Have No Time: Daily Rhythms and the Organization of Time for Low-Income Families." *Family Relations* 53: 168-78.
- Roy, M., D. Millimet, and R. Tchernis. (2012). "Federal Nutrition Programs and Childhood Obesity: Inside the Black Box." *Review of Economics of the Household* 10: 1-38.
- Ruhm, C. (2004). Parental Employment and Child Cognitive Development. *Journal of Human Resources*, 39(1), 155-192.
- Ruhm, C. (2008). Maternal Employment and Adolescent Development. *Labour Economics*, 15, 958-983.
- Schreier, H.M.C., and E. Chen. (2010). "Longitudinal Relationships Between Family Routines and Biological Profiles Among Youth With Asthma." *Health Psychology* 29: 82-90.
- Shemilt, I., M O'Brien, J. Thoburn, I. Harvey, P. Belderson, J. Robinson, and M. Camina. (2003). "School Breakfast Clubs, Children and Family Support." *Children and Society* 17: 100-12.
- Skaggs, D., Lehmann, C., Rice, C., Killelea, B., Bauer, R., Kay, R., and Vitale, M. (2005) Access to Orthopedic Care for Children with Medicaid versus Private Insurance: Results of a National Survey. *Journal of Pediatric Orthopedics*, 26(3), 400-404.
- Snow, C.E., and D.E. Beals. (2006). "Mealtime Talk That Supports Literacy Development." *New Directions for Child and Adolescent Development* 111: 51-66.
- Strauss J., Thomas D. (2007). Health over the life course. *Handbook of Development Economics*, Volume 4, Elsevier.

- Terza J., Basu A., and Rathouz P. (2008). Two-Stage Residual Inclusion Estimation: Addressing Endogeneity in Health Econometric Modeling. *Journal of Health Economics*, 27, 531–543.
- Terza, J.V. (2011). Correct Standard Errors for Multi-Stage Causal Effect Estimation in the Context of Empirical Health Policy Analysis: A Practitioner's Guide. Unpublished Manuscript, University of North Carolina at Greensboro.
- Terza, J.V. (2012). Correct Standard Errors for Multi-Stage Regression-Based Estimators: A Practitioner's Guide with Illustrations. Unpublished Manuscript, University of North Carolina at Greensboro.
- Todd, P. E., and Wolpin, K. I. (2007). The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps. *Journal of Human Capital*, 1(1), 91-136.
- Tubbs, C.Y., K.M. Roy, L.M. Burton. (2005). "Family Ties: Constructing Family Time in Low-Income Families." *Family Process* 44: 77-91.
- U.S. Food and Nutrition Service. (2010). *Nutrition Education and Promotion: The Role of FNS in Helping Low-Income Families Make Healthier Eating and Lifestyle Choices. A Report to Congress*. Alexandria, VA: U.S. Food and Nutrition Service, USDA.
- Vickery, C. (1977). "The Time-Poor: A New Look at Poverty." *Journal of Human Resources* 12: 27-48.
- Waehrer, Geetha. (2008). "The School Breakfast Program and Breakfast Consumption." Discussion Paper no. 1360-08. Madison, WI: Institute for Research on Poverty.
- Weissman J. and Bigby J. (2009). Massachusetts Health Care Reform — Near-Universal Coverage at What Cost? *New England Journal of Medicine*, 361(21), 2012-2015.
- Wells, J. (1999). Natural Selection and Sex Differences in Morbidity and Mortality in Early Life. *Journal of Theoretical Biology*, 202, 65-76.

Woodward, J.V. (2012). "Time as an Ingredient in Meal Production and Consumption." Unpublished manuscript. Greensboro, NC: University of North Carolina at Greensboro.

Woodward, J.V., and D.C. Ribar. (2012). "Is Individually-Targeted Food Assistance Shared among Family Members?" *Food and Nutrition Sciences* 3: 747-59.

Yelowitz A. and Cannon M. (2010). The Massachusetts Health Plan Much Pain, Little Gain. *Policy Analysis*, 657.

Ziebarth N. (2010). Measurement of Health, Health Inequality, and Reporting Heterogeneity. *Social Science and Medicine*, 71, 116-124.

APPENDIX A

ESTIMATING EFFECT ON HEALTH USING SERIES OF PROBITS

Dependent Variable:	P(Fair or Better)	P(Good or Better)	P(Very Good or Better)	P(Excellent)
Coefficient Estimates				
MA*During	0.040 (0.010)***	0.030 (0.007)***	0.028 (0.009)***	0.001 (0.007)
MA*After	0.070 (0.006)***	0.062 (0.009)***	0.056 (0.008)***	0.018 (0.007)**
Average Treatment Effect on Treated (After Period)				
MA*After	0.004 (0.0004)***	0.010 (0.001)***	0.018 (0.003)***	0.006 (0.002)**
Effect in Std. Deviations	0.018	0.027	0.037	0.014
Observations	2,879,296	2,879,296	2,879,296	2,879,296

Notes: Standard errors, heteroskedasticity-robust and clustered by state, are in parentheses. *** indicates statistically significant at the 0.1% level; ** 1% level; * 5 % level. All regressions include the individual-level control variables, state fixed effects, and fixed effects for each month in each year. Observations are weighted using the BRFSS sampling weights.

APPENDIX B

ESTIMATING EFFECT ON HEALTH USING SERIES OF LINEAR PROBABILITY MODELS

Dependent Variable:	P(Fair or Better)	P(Good or Better)	P(Very Good or Better)	P(Excellent)
Coefficient Estimates = Average Treatment Effects				
MA*During	0.002 (0.0008)*	0.007 (0.001)***	0.011 (0.003)***	-0.001 (0.002)
MA*After	0.004 (0.0004)***	0.011 (0.002)***	0.020 (0.003)***	0.003 (0.002)
Effect in Std. Deviations	0.020	0.031	0.040	0.008
Observations	2,879,296	2,879,296	2,879,296	2,879,296

Notes: Standard errors, heteroskedasticity-robust and clustered by state, are in parentheses. *** indicates statistically significant at the 0.1% level; ** 1% level; * 5 % level. All regressions include the individual-level control variables, state fixed effects, and fixed effects for each month in each year. Observations are weighted using the BRFSS sampling weights.

APPENDIX C

FALSIFICATION TESTS USING PRE-TREATMENT DATA

	Dependent Variable: Overall Health		
	Before: 2001-2003 After: 2004-2005	Before: 2001-2002 After: 2003-2005	Before: 2001-2002 During: 2003 After: 2004-2005
MA*During	--	--	-0.008 (0.008)
MA*After	-0.004 (0.003)	-0.006 (0.006)	-0.008 (0.005)
Observations	1,144,440	1,144,440	1,144,440

Notes: Coefficient estimates are shown; average treatment effects on the treated are available upon request. Standard errors, heteroskedasticity-robust and clustered by state, are in parentheses. *** indicates statistically significant at the 0.1% level; ** 1% level; * 5% level. All regressions include the individual-level control variables, state fixed effects, and fixed effects for each month in each year. The control group consists of all 50 other states. Observations are weighted using the BRFSS sampling weights.

APPENDIX D

CORRELATIONS BETWEEN OVERALL HEALTH AND OTHER HEALTH OUTCOMES

	Correlation with Overall Health
Days not in Good Physical Health	-0.472***
Days not in Good Mental Health	-0.255***
Days with Health Limitations	-0.381***
Activity-Limiting Joint Pain	-0.322***
BMI	-0.232***
Minutes of Moderate Exercise	0.063***
Minutes of Vigorous Exercise	0.130***
Smoker	-0.118***

*** indicates statistically significant at the 0.1% level; ** 1% level; * 5 % level. Observations are weighted using the BRFSS sampling weights.

APPENDIX E

INSTRUMENTAL VARIABLES: STRATIFIED BY GENDER AND AGE

	Gender		Age			
	Women	Men	18-34	35-44	45-54	55-64
First Stage: Any Insurance Coverage						
Coefficient Estimates						
MA*During	0.023 (0.002)***	0.016 (0.003)***	0.027 (0.004)***	0.016 (0.004)***	0.012 (0.002)***	0.021 (0.002)***
MA*After	0.042 (0.003)***	0.066 (0.004)***	0.081 (0.006)***	0.048 (0.003)***	0.043 (0.003)***	0.042 (0.002)***
1 st Stage F Statistic	142.34	201.3	105.63	159.19	138.04	224.35
Second Stage: Overall Health						
Coefficient Estimates						
Insurance	1.114 (0.159)***	0.420 (0.107)***	0.355 (0.122)**	0.413 (0.164)*	0.871 (0.164)***	1.424 (0.266)***
1 st Stage Residual	-1.075 (0.156)***	-0.402 (0.109)***	-0.209 (-0.126)	-0.401 (0.162)*	-0.977 (0.165)***	-1.566 (0.275)***
Local Average Treatment Effects						
P(Poor)	-0.135 (0.019)***	-0.030 (0.009)***	-0.011 (0.003)***	-0.025 (0.010)*	-0.111 (0.017)***	-0.302 (0.046)***
P(Fair)	-0.156 (0.017)***	-0.058 (0.015)***	-0.047 (0.014)***	-0.056 (0.022)*	-0.122 (0.020)***	-0.142 (0.020)***
P(Good)	-0.084 (0.015)***	-0.063 (0.015)***	-0.073 (0.027)**	-0.068 (0.028)*	-0.067 (0.020)***	-0.037 (0.015)*
P(Very Good)	0.151 (0.012)***	0.046 (0.013)***	0.029 (0.007)***	0.042 (0.017)*	0.127 (0.015)***	0.209 (0.019)***

P(Excellent)	0.223 (0.037)***	0.106 (0.026)***	0.102 (0.038)**	0.108 (0.044)*	0.172 (0.041)***	0.198 (0.053)***
Overall Effect in Std. Dev.	0.943	0.357	0.317	0.458	0.747	1.049
Observations	1,299,806	872,991	483,775	512,155	613,948	562,919

See notes for Table 2.1.

APPENDIX F

INSTRUMENTAL VARIABLES: STRATIFIED BY AGE AND INCOME

	Race				Household Income		
	White	Black	Hispanic	Other	<\$25,000	\$25k- \$75k	>\$75,000
First Stage: Any Insurance Coverage							
Coefficient Estimates							
MA* During	0.014 (0.002)***	-0.015 (0.006)**	0.041 (0.007)** *	0.074 (0.019)***	0.064 (0.008)***	0.016 (0.003)***	0.006 (0.002)**
MA* After	0.041 (0.002)***	0.056 (0.004)***	0.093 (0.006)** *	0.136 (0.011)***	0.138 (0.009)***	0.070 (0.003)***	0.015 (0.002)***
1 st Stage F Stat.	177.77	109.23	125.02	88.78	154.01	412.98	31.83
Second Stage: Overall Health							
Coefficient Estimates							
Insura nce	0.905 (0.132)***	1.31 (0.200)***	0.319 (0.157)*	0.793 (0.194)***	0.446 (0.069)***	0.405 (0.116)***	1.440 (0.538)**
1 st Stage Res.	-0.905 (0.134)***	-1.276 (0.205)***	-0.240 (-0.152)	-0.851 (0.196)***	-0.481 (0.075)***	-0.341 (0.117)***	-1.300 (0.533)*
Local Average Treatment Effects							
P(Poor)	-0.093 (0.016)***	-0.194 (0.032)***	-0.025 (0.014)	-0.114 (0.027)***	-0.068 (0.010)***	-0.026 (0.006)***	-0.138 (0.030)***
P(Fair)	-0.114 (0.014)***	-0.175 (0.015)***	-0.058 (0.029)*	-0.107 (0.024)***	-0.071 (0.011)***	-0.057 (0.015)***	-0.189 (0.043)***
P(Good)	-0.112 (0.015)***	-0.022 (0.013)*	-0.029 (0.013)*	-0.046 (0.020)*	-0.007 (0.005)	-0.069 (0.022)***	-0.182 (0.092)*
P(Very Good)	0.117 (0.014)***	0.158 (0.010)***	0.038 (-0.020)	0.106 (0.020)***	0.06 (0.008)***	0.052 (0.011)***	0.203 (0.013)***
P(Exce llent)	0.202 (0.030)***	0.233 (0.040)***	0.075 (0.035)*	0.160 (0.045)***	0.086 (0.015)***	0.101 (0.032)***	0.306 (0.165)
Overall Effect in S.D.	0.783	1.087	0.271	0.667	0.388	0.381	1.486
Obs.	1,704,544	182,200	247,267	38,786	524,090	1,036,339	612,368

See notes for Table 2.1.

APPENDIX G

ESTIMATES FROM ORDINARY LEAST SQUARES MODELS. FULL RESULTS

VARIABLES	Family routines	Breakfast routines	Dinner routines	Homework routines	Bedtime routines	Family time routines	Total hours apart	Start time first separation
School Breakfast Program	0.070 (0.120)	-0.119 (0.203)	0.146 (0.163)	0.068 (0.180)	0.004 (0.191)	0.250 (0.180)	-0.055 (1.557)	0.090 (0.693)
National School Lunch Program	-0.172 (0.158)	-0.448 (0.273)	-0.271 (0.190)	-0.262 (0.243)	0.154 (0.174)	-0.121 (0.222)	1.525 (1.347)	0.055 (1.009)
SNAP	-0.053 (0.088)	0.113 (0.142)	-0.186 (0.170)	-0.119 (0.121)	0.104 (0.148)	-0.243* (0.124)	-0.263 (1.023)	1.353** (0.684)
WIC	0.157 (0.098)	-0.311** (0.136)	0.492*** (0.172)	0.387** (0.151)	0.196 (0.171)	0.254* (0.144)	0.969 (0.966)	-0.081 (0.612)
Food pantry	-0.013 (0.143)	0.245 (0.210)	0.080 (0.196)	0.128 (0.203)	-0.159 (0.187)	-0.220 (0.165)	-0.123 (0.897)	-1.400* (0.728)
Household characteristics								
TANF participation	-0.040 (0.105)	-0.011 (0.174)	0.025 (0.178)	0.024 (0.152)	-0.136 (0.153)	-0.037 (0.143)	-0.503 (1.057)	-0.859 (0.802)
Monthly income (\$1000s)	-0.085** (0.035)	-0.079 (0.055)	-0.106** (0.052)	0.005 (0.064)	-0.077* (0.045)	-0.080 (0.059)	0.244 (0.362)	0.470 (0.321)
Own home	-0.004 (0.097)	0.245** (0.124)	-0.061 (0.160)	-0.050 (0.134)	0.090 (0.141)	-0.291* (0.160)	-1.420* (0.842)	0.344 (0.623)
Own car	-0.253*** (0.088)	-0.277** (0.119)	-0.311** (0.138)	-0.157 (0.132)	-0.325** (0.133)	-0.101 (0.122)	0.147 (0.857)	-0.922 (0.591)
Has financial accounts	0.197**	0.088	0.464***	0.146	0.190	0.048	0.877	-0.368

	(0.090)	(0.116)	(0.134)	(0.129)	(0.137)	(0.126)	(0.954)	(0.650)
Has outstanding loans	-0.043	-0.186*	-0.092	-0.146	0.130	-0.023	1.821**	-1.195*
	(0.082)	(0.109)	(0.122)	(0.135)	(0.140)	(0.132)	(0.862)	(0.661)
Number of adults	-0.030	-0.041	-0.008	-0.038	-0.047	-0.024	0.645*	0.127
	(0.046)	(0.062)	(0.073)	(0.065)	(0.071)	(0.054)	(0.388)	(0.198)
Number of minors	0.037	0.026	0.072	-0.050	0.094**	-0.042	-0.183	-0.312*
	(0.031)	(0.053)	(0.056)	(0.044)	(0.043)	(0.042)	(0.261)	(0.187)
Age of the youngest child	0.008	-0.038**	0.032*	0.011	0.021	0.018	-0.083	-0.075
	(0.012)	(0.017)	(0.019)	(0.014)	(0.018)	(0.017)	(0.102)	(0.072)
Boston	-0.129	0.062	-0.389**	-0.014	-0.077	-0.111	-0.547	0.349
	(0.095)	(0.147)	(0.159)	(0.144)	(0.131)	(0.146)	(0.969)	(0.776)
Chicago	0.076	0.159	-0.088	0.199	0.040	0.196	-1.316	1.516***
	(0.092)	(0.131)	(0.149)	(0.145)	(0.139)	(0.141)	(0.840)	(0.565)
Wave 2	-0.040	-0.011	0.025	0.024	-0.136	-0.037	-0.503	-0.859
	(0.105)	(0.174)	(0.178)	(0.152)	(0.153)	(0.143)	(1.057)	(0.802)
Caregiver characteristics								
Works full time	0.044	0.216	-0.098	-0.013	0.060	-0.001	1.443*	0.849
	(0.073)	(0.139)	(0.125)	(0.139)	(0.119)	(0.122)	(0.783)	(0.602)
Works part time	-0.156	-0.113	-0.177	-0.007	-0.152	-0.180	-0.297	0.276
	(0.096)	(0.150)	(0.155)	(0.131)	(0.143)	(0.134)	(1.047)	(0.555)
High school graduate	0.006	-0.002	0.043	-0.014	-0.032	0.017	1.649*	1.168*
	(0.097)	(0.144)	(0.150)	(0.133)	(0.131)	(0.127)	(0.873)	(0.678)
College education	0.135	0.125	0.077	0.198	0.097	0.239*	-0.678	0.189
	(0.098)	(0.135)	(0.157)	(0.139)	(0.151)	(0.135)	(1.057)	(0.906)
Married spouse present	0.234**	0.225*	0.106	0.173	0.369**	0.234*	-3.072***	1.423
	(0.102)	(0.135)	(0.153)	(0.145)	(0.143)	(0.142)	(1.006)	(0.880)
Cohabiting	-0.095	-0.364*	-0.177	-0.069	0.161	-0.001	-0.086	0.839
	(0.103)	(0.216)	(0.156)	(0.143)	(0.171)	(0.134)	(1.231)	(0.882)
Foreign born	0.085	0.198	0.107	0.274*	0.208	-0.175	0.299	-0.025
	(0.115)	(0.167)	(0.161)	(0.158)	(0.158)	(0.162)	(0.879)	(0.586)

Health status	-0.010 (0.029)	0.023 (0.058)	-0.038 (0.048)	0.032 (0.058)	-0.025 (0.046)	-0.001 (0.052)	0.074 (0.312)	-0.355 (0.296)
Disability status	0.050 (0.091)	0.102 (0.145)	0.095 (0.127)	0.050 (0.142)	-0.088 (0.126)	0.090 (0.134)	0.510 (0.807)	0.145 (0.532)
Black	-0.162 (0.114)	-0.035 (0.170)	-0.345* (0.181)	-0.169 (0.183)	-0.196 (0.170)	-0.074 (0.162)	-0.013 (1.184)	0.589 (0.896)
Hispanic	-0.026 (0.129)	0.112 (0.178)	-0.183 (0.196)	-0.124 (0.196)	-0.098 (0.184)	0.066 (0.184)	-0.043 (1.179)	-0.362 (0.849)
Age	-0.007 (0.006)	-0.007 (0.008)	-0.021*** (0.008)	-0.004 (0.009)	-0.009 (0.008)	0.009 (0.007)	0.046 (0.047)	0.020 (0.040)
People who will listen	0.068 (0.083)	0.005 (0.142)	0.144 (0.142)	0.108 (0.132)	-0.172 (0.117)	0.296** (0.115)	-0.593 (0.835)	0.308 (0.758)
People who help with childcare	-0.030 (0.076)	0.108 (0.140)	-0.165 (0.114)	-0.068 (0.153)	0.229** (0.112)	-0.293** (0.114)	0.480 (0.853)	0.001 (0.701)
People who help with small favors	0.133 (0.089)	-0.145 (0.127)	0.134 (0.147)	0.246 (0.170)	0.264** (0.128)	0.280* (0.151)	-0.126 (1.056)	-2.068** (1.043)
People who will loan money	0.192** (0.082)	0.122 (0.136)	0.351*** (0.128)	-0.023 (0.149)	0.050 (0.116)	0.245* (0.137)	0.357 (0.906)	0.899 (0.681)
Child characteristics								
Female	-0.177*** (0.066)	-0.145 (0.096)	-0.151 (0.109)	-0.129 (0.105)	-0.082 (0.110)	-0.330*** (0.086)	-0.001 (0.640)	0.362 (0.494)
Number of observations	1,006	1,006	1,006	1,006	1,006	1,006	614	512

Note: Estimates from authors' linear regressions using school-enrolled children living with a continuing caregiver interviewed in the first and second waves of the Three-City Study. Regressions also control for age of the focal and month of the interview. Estimates incorporate sampling weights.

Clustered standard errors are shown in parentheses.

***Coefficient different from zero $p < 0.01$ **Coefficient different from zero $p < 0.05$ *Coefficient different from zero $p < 0.1$

APPENDIX H

ESTIMATES FROM FIXED EFFECTS MODELS. FULL RESULTS

VARIABLES	Family routines	Breakfast routines	Dinner routines	Homework routines	Bedtime routines	Family time routines	Total hours apart	Start time first separation
School breakfast program	0.162 (0.162)	0.177 (0.201)	0.008 (0.219)	0.006 (0.248)	0.174 (0.275)	0.291 (0.300)	1.343 (1.741)	-2.381** (1.051)
National lunch school program	0.012 (0.165)	-0.255 (0.273)	-0.025 (0.207)	-0.005 (0.360)	0.125 (0.214)	0.201 (0.299)	1.734 (1.961)	-1.843 (1.518)
SNAP	-0.021 (0.134)	0.200 (0.190)	-0.323 (0.208)	-0.324* (0.166)	-0.032 (0.194)	0.069 (0.271)	-0.737 (1.336)	0.079 (0.864)
WIC	0.321* (0.166)	-0.248 (0.224)	0.588** (0.257)	0.773*** (0.251)	0.464 (0.324)	0.481** (0.205)	1.102 (1.211)	-1.196 (0.998)
Food pantry	-0.018 (0.138)	0.104 (0.212)	0.193 (0.262)	0.146 (0.217)	0.003 (0.174)	-0.372* (0.191)	0.068 (1.327)	-3.778*** (1.045)
Household characteristics								
TANF participation	-0.058 (0.165)	-0.033 (0.288)	-0.084 (0.229)	0.134 (0.238)	0.082 (0.189)	-0.197 (0.252)	-5.262*** (1.456)	2.087** (0.835)
Monthly income (\$1000s)	-0.080* (0.047)	-0.099 (0.068)	-0.117* (0.066)	0.002 (0.078)	-0.029 (0.068)	-0.074 (0.084)	1.227** (0.483)	-0.055 (0.396)
Own home	-0.128 (0.144)	0.213 (0.224)	-0.073 (0.172)	-0.215 (0.217)	-0.091 (0.276)	-0.560* (0.292)	2.132 (1.459)	0.531 (1.250)
Own car	-0.188 (0.127)	-0.487*** (0.186)	-0.227 (0.217)	0.007 (0.213)	-0.189 (0.187)	0.152 (0.205)	-1.413 (1.485)	1.060 (1.352)
Has financial accounts	-0.240** (0.110)	-0.101 (0.159)	-0.266* (0.157)	-0.079 (0.166)	-0.540*** (0.201)	-0.054 (0.188)	2.089 (1.289)	-0.077 (0.962)
Has outstanding loans	-0.003 (0.132)	-0.151 (0.151)	0.057 (0.197)	0.065 (0.174)	-0.022 (0.183)	0.104 (0.217)	-0.031 (1.038)	-0.881 (0.752)
Number of adults	0.011 (0.071)	0.067 (0.118)	0.066 (0.087)	-0.189* (0.097)	-0.122 (0.100)	0.034 (0.118)	-0.062 (0.721)	0.232 (0.451)

Number of minors	0.156** (0.075)	0.308*** (0.112)	0.151 (0.106)	0.317** (0.141)	0.092 (0.082)	0.073 (0.113)	0.518 (0.615)	-1.081** (0.442)
Age of the youngest child	0.041* (0.024)	0.011 (0.036)	0.058* (0.033)	0.037 (0.036)	0.072** (0.029)	0.022 (0.037)	-0.125 (0.196)	-0.326** (0.160)
Wave 2	-0.164** (0.075)	-0.203 (0.142)	-0.191* (0.114)	-0.042 (0.127)	-0.132 (0.131)	-0.129 (0.119)	-0.572 (0.888)	-1.756*** (0.534)
Caregiver characteristics								
Works full time	0.049 (0.136)	0.401* (0.216)	-0.068 (0.199)	-0.383* (0.232)	0.102 (0.208)	-0.238 (0.215)	-0.733 (1.395)	0.915 (1.201)
Works part time	-0.128 (0.142)	0.124 (0.216)	-0.060 (0.201)	-0.448** (0.203)	-0.208 (0.197)	-0.368 (0.229)	-2.312 (1.642)	2.454** (1.198)
Married spouse present	0.190 (0.153)	0.206 (0.221)	0.113 (0.233)	0.125 (0.260)	0.213 (0.199)	0.229 (0.273)	-5.103*** (1.843)	3.817*** (1.330)
Cohabiting	0.256 (0.169)	0.422 (0.386)	0.325 (0.268)	0.664** (0.316)	0.509* (0.270)	-0.232 (0.302)	-0.061 (1.738)	0.230 (1.164)
Health status	0.047 (0.051)	0.019 (0.081)	0.127* (0.072)	0.029 (0.082)	0.011 (0.066)	0.032 (0.083)	0.816* (0.490)	-0.689** (0.301)
Disability status	0.190* (0.114)	-0.101 (0.170)	0.411*** (0.158)	0.273 (0.199)	0.263 (0.172)	0.188 (0.215)	-0.456 (1.517)	0.851 (0.951)
People who will listen	0.107 (0.108)	0.310* (0.160)	-0.012 (0.136)	0.046 (0.154)	-0.015 (0.185)	0.145 (0.167)	-0.513 (0.944)	0.068 (0.865)
People who help with Childcare	0.029 (0.101)	-0.037 (0.134)	0.016 (0.124)	-0.128 (0.184)	0.241 (0.160)	-0.105 (0.147)	0.377 (1.012)	0.666 (0.854)
People who help with small favors	0.069 (0.092)	-0.395*** (0.146)	0.295** (0.149)	-0.034 (0.252)	0.200 (0.166)	0.177 (0.179)	-1.737 (1.538)	-2.578** (1.272)
People who will loan Money	0.186 (0.114)	0.327** (0.156)	0.152 (0.157)	0.264 (0.199)	0.037 (0.171)	0.228 (0.181)	0.654 (1.432)	0.654 (0.981)
Number of observations	1,006	1,006	1,006	1,006	1,006	1,006	614	512
Number of families	503	503	503	503	503	503	307	256

Note: Estimates from authors' fixed effects regressions using school-enrolled children living with a continuing caregiver interviewed in the first and second waves of the Three-City Study. Regressions also control for month of the interview. Estimates incorporate sampling weights. Robust standard errors are shown in parentheses. ***Coefficient different from zero $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.